534 Rec'd PCT/PTO 05 JAN 2001

# DESCRIPTION

## Human Proteins Having Hydrophobic Domains and DNAs Encoding These Proteins

### TECHNICAL FIELD

The present invention relates to human proteins having hydrophobic domains, DNAs coding for these proteins, and expression vectors for these DNAs as well as eucaryotic cells expressing these DNAs. The proteins of the present invention can be employed as pharmaceuticals or as antigens for preparing antibodies against these proteins. The human cDNAs of the present invention can be utilized as probes for the genetic diagnosis and gene sources for the gene therapy. Furthermore, the cDNAs can be utilized as gene sources for large-scale production of the proteins encoded by these cDNAs. Cells into which these genes are introduced to express secretory proteins and membrane proteins in large amounts can be utilized for detection of the corresponding receptors and ligands, screening of novel low-molecular pharmaceuticals, and so on.

### BACKGROUND ART

Cells secrete many proteins outside the cells. These secretory proteins play important roles for the proliferation control, the differentiation induction, the material transportation, the biological protection, etc. in the cells. Different from intracellular proteins, the secretory proteins exert their actions outside the cells,

hidden potentialities as medicines. In fact, a number of human secretory proteins such as interferons, interleukins, erythropoietin, thrombolytic agents, etc. have been currently employed as medicines. In addition, secretory proteins other than those described above have been undergoing clinical trials to develop as pharmaceuticals. Because it has been conceived that the human cells still produce many unknown secretory proteins, availability of these secretory proteins as well as genes coding for them is expected to lead to development of novel pharmaceuticals utilizing these proteins.

On the other hand, membrane proteins play important roles, as signal receptors, ion channels, transporters, etc. in the material transportation and the information transmission through the cell membrane. Examples thereof include receptors for a variety of cytokines, ion channels for the sodium ion, the potassium ion, the chloride ion, etc., transporters for saccharides and amino acids, and so on, where the genes for many of them have been cloned already. It has been clarified that abnormalities of these membrane proteins are associated with a number of hitherto-cryptogenic diseases. Therefore, discovery of a new membrane protein is anticipated to lead to elucidation of the causes of many diseases, so that isolation of a new gene coding for the membrane protein has been desired.

Heretofore, owing to difficulty in the purification from human cells, these secretory proteins and membrane proteins have been isolated by an approach from the gene side. A general method is the so-called expression cloning [illegible] eucaryotic cells to express cDNAs and then screening of the cells secreting, or expressing on the surface of membrane,

the objective active protein. However, this method is applicable only to cloning of a gene for a protein with a known function.

In general, secretory proteins and membrane proteins possess at least one hydrophobic domain inside the proteins, wherein, after synthesis thereof in the ribosome, this domain works as a secretory signal or remains in the phospholipid membrane to be trapped in the membrane. Accordingly, the evidence of this cDNA for encoding a secretory protein and a membrane protein is provided by determination of the whole base sequence of a full-length cDNA followed by detection of highly hydrophobic domain(s) in the amino acid sequence of the protein encoded by this cDNA.

## OBJECTS OF THE INVENTION

The main object of the present invention is to provide novel human proteins having hydrophobic domains, DNAs coding for these proteins, and expression vectors for these DNAs as well as transformed eucaryotic cells that are capable of expressing these DNAs. This object as well as other objects and advantages of the present invention will become apparent to those skilled in the art from the following description with reference to the accompanying drawings.

## BRIEF DESCRIPTION OF DRAWINGS

Fig. 1 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP01550.

Fig. 2 illustrates the hydrophobicity/hydrophilicity [illegible]

Fig. [illegible] illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10195.

Fig. 4 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10423.

Fig. 5 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10506.

Fig. 6 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10507.

Fig. 7 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10548.

Fig. 8 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10566.

Fig. 9 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10567.

Fig. 10 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10568.

Fig. 11 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP01426.

Fig. 12 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP02515.

Fig. 13 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP02575.

Fig. 14 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10357.

Fig. 15 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10447.

Fig. 16 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10477.

Fig. 17 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10513.

Fig. 18 illustrates the hydrophobicity/hydrophilicity [illegible]

Fig. 19 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10557.

Fig. 20 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10563.

Fig. 21 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP01467.

Fig. 22 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP01956.

Fig. 23 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP02545.

Fig. 24 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP02551.

Fig. 25 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP02631.

Fig. 26 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP02632.

Fig. 27 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10488.

Fig. 28 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10538.

Fig. 29 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10542.

Fig. 30 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10571.

Fig. 31 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP01470.

Fig. 32 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP02419.

Fig. 33 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP02631.

Fig. 34 illustrates the hydrophobicity/hydrophilicity [illegible]

Fig. 35 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10031.

Fig. 36 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10530.

Fig. 37 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10541.

Fig. 38 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10550.

Fig. 39 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10590.

Fig. 40 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10591.

Fig. 41 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP01462.

Fig. 42 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP02485.

Fig. 43 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP02798.

Fig. 44 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10041.

Fig. 45 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10246.

Fig. 46 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10392.

Fig. 47 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10489.

Fig. 48 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10519.

Fig. 49 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10531.

Fig. 50 illustrates the hydrophobicity/hydrophilicity [illegible]

SUMMARY OF THE INVENTION

As the result of intensive studies, the present inventors have been successful in cloning of cDNAs coding for proteins having hydrophobic domains from the human full-length cDNA bank, thereby completing the present invention. In other words, the present invention provides human proteins having hydrophobic domains, namely proteins comprising any of the amino acid sequences represented by SEQ ID Nos. 1 to 10, 31 to 40, 61 to 70, 91 to 100, and 121 to 130. Moreover, the present invention provides DNAs coding for the above-mentioned proteins, exemplified by cDNAs comprising any of the base sequences represented by SEQ ID Nos. 11 to 20, 41 to 50, 71 to 80, 101 to 110, and 131 to 140, as well as expression vectors that are capable of expressing any of these DNAs by in vitro translation or in eucaryotic cells and transformed eucaryotic cells that are capable of expressing these DNAs and of producing the above-mentioned proteins.

DETAILED DESCRIPTION OF THE INVENTION

The proteins of the present invention can be obtained, for example, by a method for isolation from human organs, cell lines, etc., a method for preparation of peptides by the chemical synthesis, or a method for production with the recombinant DNA technology using the DNAs coding for the hydrophobic domains of the present invention, among which the method for production with the recombinant DNA technology is employed preferably. For instance, in vitro expression of the proteins can be achieved by preparation of an RNA by in vitro transcription from a vector having one of [illegible] translation using this RNA as a template. Also, introduction of the translated region into a suitable expression vector

by the method known in the art leads to expression of a large amount of the encoded protein in prokaryotic cells such as *Escherichia coli, Bacillus subtilis*, etc., and eucaryotic cells such as yeasts, insect cells, mammalian cells, etc.

In the case where one of the proteins of the present invention is produced by expressing the DNA by in vitro translation, the protein of the present invention can be produced in vitro, when the translated region of this cDNA is introduced into a vector having an RNA polymerase promoter, followed by addition of the vector to an in vitro translation system such as a rabbit reticulocyte lysate or a wheat germ extract, containing an RNA polymerase corresponding to the promoter. RNA polymerase promoters are exemplified by T7, T3, SP6, and the like. The vectors containing these RNA polymerase promoters are exemplified by pKA1, pCDM8, pT3/T7 18, pT7/3 19, pBluescript II, and so on. Furthermore, the protein of the present invention can be expressed as the secreted form or the form incorporated into the microsome membrane, when a canine pancreas microsome or the like is added to the reaction system.

In the case where one of the protein of the present invention is produced by expressing the DNA in a microorganism such as *Escherichia coli* etc., a recombinant expression vector bearing the translated region of the cDNA of the present invention is constructed in an expression vector having an origin which can be replicated in the microorganism, a promoter, a ribosome-binding site, a cDNA-cloning site, a terminator etc. and, after transformation of [illegible] transformant is incubated, whereby the protein encoded by said cDNA can be produced on a large scale in the

microorganism. In this case, a protein fragment containing any region can be obtained by carrying out the expression with inserting an initiation codon and a termination codon in front of and behind the selected translated region. Alternatively, a fusion protein with another protein can be expressed. Only the portion of the protein encoded by this cDNA can be obtained by cleavage of this fusion protein with a suitable protease. The expression vector for *Escherichia coli* is exemplified by the pUC series, pBluescript II, the pET expression system, the pGEX expression system, and so on.

In the case where one of the proteins of the present invention is produced by expressing the DNA in eucaryotic cells, the protein of the present invention can be produced as a secretory protein or as a membrane protein on the cell-membrane surface, when the translated region of this cDNA is introduced into an expression vector for eucaryotic cells that has a promoter, a splicing region, a poly(A) addition site, etc., followed by introduction into the eucaryotic cells. The expression vector is exemplified by pKA1, pED6dpc2, pCDM8, pSVK3, pMSG, pSVL, pBK-CMV, pBK-RSV, EBV vector, pRS, pYES2, and so on. Examples of eucaryotic cells to be used in general include mammalian cultured cells such as simian kidney cells COS7, Chinese hamster ovary cells CHO, etc., budding yeasts, fission yeasts, silkworm cells, Xenopus oocytes, and so on, but any eucaryotic cells may be used, provided that they are capable of expressing the proteins of the present invention. The expression vector can be introduced into the eucaryotic cells by methods known in the art such as the electroporation method, the calcium [illegible]

method, and so on.

After one of the proteins of the present invention is

expressed in prokaryotic cells or eucaryotic cells, the objective protein can be isolated from the culture and purified by a combination of separation procedures known in the art. Such examples include treatment with a denaturing agent such as urea or a detergent, sonication, enzymatic digestion, salting-out or solvent precipitation, dialysis, centrifugation, ultrafiltration, gel filtration, SDS-PAGE, isoelectric focusing, ion-exchange chromatography, hydrophobic chromatography, affinity chromatography, reverse phase chromatography, and so on.

The proteins of the present invention include peptide fragments (5 amino acid residues or more) containing any partial amino acid sequence in the amino acid sequences represented by SEQ ID Nos. 1. to 10, 31 to 40, 61 to 70, 91 to 100, and 121 to 130. These peptide fragments can be utilized as antigens for preparation of antibodies. Hereupon, among the proteins of the present invention, those having the signal sequences are secreted in the form of mature proteins, after the signal sequences are removed. Therefore, these mature proteins shall come within the scope of the present invention. The N-terminal amino acid sequences of the mature proteins can be easily determined by using the method for the determination of cleavage site of a signal sequence [JP 8-187100 A]. Furthermore, some membrane proteins undergo the processing on the cell surface to be converted to the secretory forms. Such proteins or peptides in the secretory forms shall come within the scope of the present invention. In the case where sugar chain-binding sites are present in the amino acid sequences, expression in [illegible] chains are attached. Accordingly, such proteins or peptides to which sugar chains are attached shall come within the

scope of the present invention.

The DNAs of the present invention include all the DNAs coding for the above-mentioned proteins. These DNAs can be obtained by using a method by chemical synthesis, a method by cDNA cloning, and so on.

The cDNAs of the present invention can be cloned, for example, from cDNA libraries derived from the human cells. These cDNAs are synthesized by using as templates poly(A)$^+$ RNAs extracted from human cells. The human cells may be cells delivered from the human body, for example, by the operation or may be the cultured cells. The cDNAs can be synthesized by using any method selected from the Okayama-Berg method [Okayama, H. and Berg, P., Mol. Cell. Biol. 2: 161-170 (1982)], the Gubler-Hoffman method [Gubler, U. and Hoffman, J. Gene 25: 263-269 (1983)], and so on, but it is preferred to use the capping method [Kato, S. et al., Gene 150: 243-250 (1994)], as exemplified in Examples, in order to obtain a full-length clone in an effective manner. In addition, commercially available, human cDNA libraries can be utilized. Cloning of the cDNAs of the present invention from the cDNA libraries can be carried out by synthesis of an oligonucleotide on the basis of base sequences of any portion in the cDNA of the present invention, followed by screening using this oligonucleotide as the probe according to the colony or plaque hybridization by a method known in the art. In addition, the cDNA fragments of the present invention can be prepared by synthesis of oligonucleotides which hybridize with both termini of the objective cDNA fragment, followed by the usage of these oligonucleotides as

from human cells.

The cDNAs of the present invention are characterized by

comprising either of the base sequences represented by SEQ ID Nos. 11 to 20, 41 to 50, 71 to 80, 101 to 110, and 131 to 140 or the base sequences represented by SEQ ID Nos. 21 to 30, 51 to 60, 81 to 90, 111 to 120, and 141 to 150. Table 1 summarizes the clone number (HP number), the cells from which the cDNA was obtained, the total base number of the cDNA, and the number of the amino acid residues of the encoded protein, for each of the cDNAs.

Table 1

| SEQ ID No. | HP number | Cells | Base number | Number of amino acid residues |
|---|---|---|---|---|
| 1, 11, 21 | HP01550 | Stomach cancer | 510 | 125 |
| 2, 12, 22 | HP02593 | Saos-2 | 697 | 131 |
| 3, 13, 23 | HP10195 | HT-1080 | 1619 | 242 |
| 4, 14, 24 | HP10423 | U-2 OS | 1066 | 264 |
| 5, 15, 25 | HP10506 | Stomach cancer | 618 | 112 |
| 6, 16, 26 | HP10507 | Stomach cancer | 1021 | 146 |
| 7, 17, 27 | HP10548 | Stomach cancer | 1432 | 344 |
| 8, 18, 28 | HP10566 | Stomach cancer | 601 | 97 |
| 9, 19, 29 | HP10567 | Stomach cancer | 585 | 124 |
| 10, 20, 30 | HP10568 | Stomach cancer | 1100 | 327 |
| 31, 41, 51 | HP01426 | Stomach cancer | 1065 | 313 |
| 32, 42, 52 | HP02515 | Saos-2 | 937 | 229 |
| 33, 43, 53 | HP02575 | Saos-2 | 1678 | 467 |
| 34, 44, 54 | HP10357 | Stomach cancer | 467 | 99 |
| 35, 45, 55 | HP10447 | Liver | 875 | 189 |
| 36, 46, 56 | HP10477 | Liver | 1256 | 363 |
| 37, 47, 57 | HP10513 | Stomach cancer | 884 | 249 |
| 38, 48, 58 | HP10540 | Saos-2 | 589 | 98 |
| 39, 49, 59 | HP10557 | Stomach cancer | 673 | 172 |
| 40, 50, 60 | HP10563 | Saos-2 | 1425 | 120 |
| 61, 71, 81 | HP01467 | HT-1080 | 1436 | 307 |
| 62, 72, 82 | HP01956 | Liver | 997 | 183 |
| 63, 73, 83 | HP02545 | Saos-2 | 1753 | 327 |
| 64, 74, 84 | HP02551 | Saos-2 | 1117 | 223 |
| 65, 75, 85 | HP02631 | Saos-2 | 1380 | 48 |
| 66, 76, 86 | HP02632 | HT-1080 | 1503 | 371 |
| [illegible] | HP10488 | Liver | 733 | 90 |
| [illegible] | HP1054[illegible] | Stomach cancer | [illegible] | [illegible] |
| 70, 80, 90 | HP10571 | Stomach cancer | 1229 | 152 |

| | | | | |
|---|---|---|---|---|
| 91, 101, 111 | HP01470 | Stomach cancer | 1619 | 358 |
| 92, 102, 112 | HP02419 | Stomach cancer | 2054 | 226 |
| 93, 103, 113 | HP02631 | Saos-2 | 1380 | 195 |
| 94, 104, 114 | HP02695 | Stomach cancer | 1292 | 339 |
| 95, 105, 115 | HP10031 | Saos-2 | 2168 | 487 |
| 96, 106, 116 | HP10530 | Saos-2 | 1357 | 393 |
| 97, 107, 117 | HP10541 | Stomach cancer | 711 | 196 |
| 98, 108, 118 | HP10550 | Stomach cancer | 651 | 107 |
| 99, 109, 119 | HP10590 | HT-1080 | 1310 | 350 |
| 100, 110, 120 | HP10591 | HT-1080 | 1400 | 107 |
| 121, 131, 141 | HP01462 | HT-1080 | 2050 | 483 |
| 122, 132, 142 | HP02485 | Stomach cancer | 2746 | 334 |
| 123, 133, 143 | HP02798 | HT-1080 | 1136 | 267 |
| 124, 134, 144 | HP10041 | Saos-2 | 619 | 106 |
| 125, 135, 145 | HP10246 | KB | 864 | 224 |
| 126, 136, 146 | HP10392 | U-2 OS | 1527 | 258 |
| 127, 137, 147 | HP10489 | Stomach cancer | 659 | 110 |
| 128, 138, 148 | HP10519 | Stomach cancer | 710 | 91 |
| 129, 139, 149 | HP10531 | Saos-2 | 2182 | 344 |
| 130, 140, 150 | HP10574 | Stomach cancer | 2773 | 428 |

Hereupon, the same clones as the cDNAs of the present invention can be easily obtained by screening of the cDNA libraries constructed from the human cell lines or human tissues utilized in the present invention by the use of an oligonucleotide probe synthesized on the basis of the cDNA base sequence described in any of SEQ ID Nos. 11 to 30, 41 to 60, 71 to 90, 101 to 120, and 131 to 150.

In general, the polymorphism due to the individual difference is frequently observed in human genes. Accordingly, any cDNA in which one or plural nucleotides are [illegible] in SEQ ID Nos. 11 to 30, 41 to 60, 71 to 90, 101 to 120, and

131 to 150 shall come within the scope of the present invention.

In a similar manner, any protein in which one or plural amino acids are inserted, deleted and/or substituted with other amino acids shall come within the scope of the present invention, as far as the protein possesses the activity of any protein having the amino acid sequences represented by SEQ ID Nos. 1 to 10, 31 to 40, 61 to 70, 91 to 100, and 121 to 130.

The cDNAs of the present invention include cDNA fragments (10 bp or more) containing any partial base sequence in the base sequences represented by SEQ ID Nos. 11 to 20, 41 to 50, 71 to 80, 101 to 110, and 131 to 140 or in the base sequences represented by SEQ ID Nos. 21 to 30, 51 to 60, 81 to 90, 111 to 120, and 141 to 150. Also, DNA fragments consisting of a sense strand and an anti-sense strand shall come within this scope. These DNA fragments can be utilized as the probes for the genetic diagnosis.

In addition to the activities and uses described above, the polynucleotides and proteins of the present invention may exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

Research Uses and Utilities

[illegible] can be used by the research community for various purposes. The polynucleotides can be used to express recombinant

protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodiesusing DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-

another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine

levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or [illegible] carbon source, use as a [illegible] such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be

administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immun[illegible] Human[illegible] Immun[illegible] 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular

Immunology 133:327-341, 1991; Bertagnolli, et al., J. Immunol. 149:3778-3783, 1992; Bowman et al., J. Immunol. 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon γ, Schreiber, R.D. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6- Nordan, R. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Acad. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; [illegible] Giannotti, J., Clark, S.C. and Turner, K.J. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp.

6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial orfungal infections, or may result from autoimmune disorders. More specifically, infectious [illegible] fungal or other [illegible] invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., malaria spp.

and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also to be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing [illegible] from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent

has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (e.g., B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by [illegible] necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or

tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function in vivo on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the [illegible] Additionally, blocking reagents may induce [illegible] could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating

autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of viral infection. In addition, systemic viral diseases such as influenza, the commoncold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a [illegible] described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the

transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (e.g., sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected ex vivo with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection in vivo.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected [illegible] cytoplasmic-domain truncated portion) of an MHC class I α chain protein and $_2$ microglobulin protein or an MHC class

II chain protein and an MHC class II chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488- [illegible] 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Bowmanet al., J.

Virology 61:1992-1998; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J.J. and Brunswick, M. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, [illegible] 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965,

1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells [illegible] indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to

stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence [illegible] limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and

Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In Culture of Hematopoietic Cells. R.I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and Briddell, R.A. In Culture of Hematopoietic Cells. R.I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In Culture of Hematopoietic Cells. R.I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, NY. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R.I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland, H.J. In Culture of Hematopoietic Cells. R.I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

Tissue Growth Activity

A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and [illegible] treatment [illegible]

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is

not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a [illegible] [illegible] tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and

in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors ex vivo for return in vivo to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager [illegible] accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head

trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among [illegible] be measured by the following methods:

[illegible] limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon);

International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium ).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin- group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime [illegible] sheep and pigs.

The activity of a protein of the invention may, among

other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing [illegible] chemotaxis.

The activity of a protein of the invention may, among

other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis)consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25: 1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153: 1762-1768, 1994.

<u>Hemostatic and Thrombolytic Activity</u>

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result,such a protein is expected to be useful in treatment of various coagulation disorders (includinghereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central ... vessels (e.g. stroke).

other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include,

without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in:Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. [illegible] Associates and Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22),

Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of ytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Tumor Inhibition Activity

immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A

protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth

Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or caricadic cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive [illegible] violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of

embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

Examples

The present invention is specifically illustrated in more detail by the following Examples, but Examples are not intended to restrict the present invention. The basic operations with regard to the recombinant DNA and the enzymatic reactions were carried out according to the literature ["Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Laboratory, 1989]. Unless otherwise stated, restrictive enzymes and a variety of modification enzymes to be used were those available from Takara Shuzo. The buffer compositions and the reaction conditions for each of the enzyme reactions were as described in the manufacturer's instructions. The cDNA synthesis was carried out according to the literature [Kato, S. et al., Gene 150: 243-250 (1994)].

(1) Selection of cDNAs Encoding Proteins Having Hydrophobic Domains

The cDNA library of fibrosarcoma cell line HT-1080 [illegible] Saos-2 (WO97/33993), the cDNA library of osteosarcoma cell line U-2 OS (WO98/21328), the cDNA library of epidermoid

carcinoma cell line KB (WO98/11217), the cDNA library of tissues of stomach cancer delivered by the operation (WO98/21328), the cDNA library of liver tissue delivered by the operation (WO98/21328), and were used for the cDNA libraries. Full-length cDNA clones were selected from respective libraries and the whole base sequences thereof were determined to construct a homo-protein cDNA bank consisting of the full-length cDNA clones. The hydrophobicity/hydrophilicity profiles were determined for the proteins encoded by the full-length cDNA clones registered in the homo-protein cDNA bank by the Kyte-Doolittle method [Kyte, J. & Doolittle, R. F., J. Mol. Biol. 157: 105-132 (1982)] to examine the presence or absence of a hydrophobic region. Any clone that has a hydrophobic region being putative as a secretory signal or a transmembrane domain in the amino acid sequence of the encoded protein was selected as a clone candidate.

(2) Protein Synthesis by In Vitro Translation

The plasmid vector bearing the cDNA of the present invention was used for in vitro transcription/translation with a $T_NT$ rabbit reticulocyte lysate kit (Promega). In this case, [$^{35}$S]methionine was added to label the expression product with a radioisotope. Each of the reactions was carried out according to the protocols attached to the kit. Two micrograms of the plasmid was subjected to the reaction at 30°C for 90 minutes in the reaction solution of a total volume of 25 μl containing 12.5 μl μ of $T_NT$ rabbit reticulocyte lysate, 0.5 μl of a buffer solution (attached [illegible] amino acid mixture (without [illegible]

0.5 μl of T7 RNA polymerase, and 20 U of RNasin. Also, an experiment in the presence of a membrane system was carried

out by adding to this reaction system 2.5 μl of a canine pancreas microsome fraction (Promega). To 3 μl of the resulting reaction solution was added 2 μl of the SDS sampling buffer (125 mM Tris-hydrochloric acid buffer, pH 6.8, 120 mM 2-mercaptoethanol, 2% SDS solution, 0.025% bromophenol blue, and 20% glycerol) and the resulting mixture was heated at 95°C for 3 minutes and then subjected to SDS-polyacrylamide gel electrophoresis. The molecular weight of the translation product was determined by carrying out the autoradiography.

(3) Expression by COS7

*Escherichia coli* cells bearing the expression vector for the protein of the present invention was incubated at 37°C for 2 hours in 2 ml of the 2xYT culture medium containing 100 μg/ml of ampicillin, the helper phage M13KO7 (50 μl) was added, and the incubation was continued at 37°C overnight. A supernatant separated by centrifugation underwent precipitation with polyethylene glycol to obtain single-stranded phage particles. These particles were suspended in 100 μl of 1 mM Tris-0.1 mM EDTA, pH 8 (TE).

The cultured cells derived from simian kidney, COS7, were incubated at 37°C in the presence of 5% $CO_2$ in the Dulbecco's modified Eagle's culture medium (DMEM) containing 10% fetal calf serum. Into a 6-well plate (Nunc, well diameter: 3 cm) were inoculated with 1 x $10^5$ COS7 cells and incubation was carried out at 37°C for 22 hours in the presence of 5% $CO_2$. After the culture medium was removed, the cell surface was washed with a phosphate buffer solution and then washed again with DMEM containing 50 mM Tris [illegible] MEM [illegible] was added a suspension of [illegible] μl of the single-stranded phage suspension, 0.6 ml of the DMEM culture medium, and 3 μl of

TRANSFECTAM™ (IBF) and the resulting mixture was incubated at 37°C for 3 hours in the presence of 5% $CO_2$. After the sample solution was removed, the cell surface was washed with TDMEM, 2 ml per well of DMEM containing 10% fetal calf serum was added, and the incubation was carried out at 37°C for 2 days in the presence of 5% $CO_2$. After the culture medium was replaced by a culture medium containing [$^{35}$S]cystine or [$^{35}$S]methionine, the incubation was carried out for one hour. After the culture medium and the cells were separated by centrifugation, proteins in the culture medium fraction and the cell-membrane fraction were subjected to SDS-PAGE.

(4) Clone Examples

<HP01550> (SEQ ID Nos. 1, 11, and 21)

Determination of the whole base sequence of the cDNA insert of clone HP01550 obtained from cDNA library of human stomach cancer revealed the structure consisting of a 65-bp 5′-untranslated region, a 378-bp ORF, and a 67-bp 3′-untranslated region. The ORF codes for a protein consisting of 125 amino acid residues and there existed one putative transmembrane domain. Figure 1 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 15 kDa that was almost identical with the molecular weight of 13,825 predicted from the ORF.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Caenorhabditis elegans [illegible] Table [illegible] shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the C.

elegans hypothetical protein F45G2.c (CE). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 44.5% in the entire region.

Table 2

```
HP  MAKYLAQIIVMGVQVVGRAFARALRQEF------AASRAAADARGRAGHRSAAASNLS-
        .. .  . ..*..*..**.*.*.      **..**..... .. *..*.** .
CE MPWRTALKVALAAGEAVAKALTRAVRDEIKQTQQAAARHAASTGQSASETRENANSNAKL
HP GLSLQEAQQILNV-SKLSPEEVQKNYEHLFKVNDKSVGGSFYLQSKVVRAKERLDEEL-K
   *.**.*. ***** . *..***.*.*****..**** **..****** *****.***. .
CE GISLEESLQILNVKTPLNREEVEKHYEHLFNINDKSKGGTLYLQSKVFRAKERIDEEFGR
HP IQAQEDREKGQMPHT
   *. .*...*.. ..*
CE IELKEEKKKEENAKTE
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA338859) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP02593> (SEQ ID Nos. 2, 12, and 22)

Determination of the whole base sequence of the cDNA [illegible] osteosarcoma cell line Saos-2 revealed the structure consisting of a 103-bp 5′-untranslated region, a 396-bp ORF,

and a 198-bp 3'-untranslated region. The ORF codes for a protein consisting of 131 amino acid residues and there existed four putative transmembrane domains at the C-terminus. Figure 2 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of a high molecular weight.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to a human OB-R gene-related protein (EMBL Accession No. Y12670). Table 3 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the human OB-R gene-related protein (OB). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 67.9% in the entire region.

Table 3

```
HP MAGIKALISLSFGGAIGLMFLMLGCALPIYNKYWPLFVLFFYILSPIPYCIARRLVDDTD
   ***.***..***.***** ********  *. *******.*. .****. **.*.. *.*
OB MAGVKALVALSFSGAIGLTFLMLGCALEDYGVYWPLFVLIFHAISPIPHFIAKRVTYDSD
HP AMSNACKELAIFLTTGIVVSAFGLPIVFARAHLIEWGACALVLTGNTVIFATILGFFLVF
   * *.**.*** *.**********.*...**. .*.****.***.**.*** ** ****.*
OB ATSSACRELAYFFTTGIVVSAFGFPVILARVAVIKWGACGLVLAGNAVIFLTIQGFFLIF
HP GSNDDFSWQQW
   *  *****  **
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA306490) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10195> (SEQ ID Nos. 3, 13, and 23)

Determination of the whole base sequence of the cDNA insert of clone HP10195 obtained from cDNA library of human fibrosarcoma HT-1080 revealed the structure consisting of a 286-bp 5'-untranslated region, a 729-bp ORF, and a 604-bp 3'-untranslated region. The ORF codes for a protein consisting of 242 amino acid residues and there existed one putative transmembrane domain at the C-terminus. Figure 3 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 32 kDa that was somewhat larger than the molecular weight of 27,300 predicted from the ORF. When expressed in COS7 cells, an expression product of about 21 kDa was observed in the supernatant fraction and the membrane fraction.

The search of the protein data base using the amino acid sequence of the present protein has revealed the registration of sequences that were similar to the Aplysia VAP-33 (SWISS-PROT Accession No. P53173). Table 4 shows the comparison between amino acid sequences of the human protein [illegible]

Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the

present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 46.5% in the entire region.

Table 4

```
HP MAKHEQILVLDPPTDLKFKGPFTDVVTTNLKLRNPSDRKVCFKVKTTAPRRYCVRPNSGI
   **.*** *.*.*...*.**********..***.**.**..*********.**********
AP MASHEQALILEPAGELRFKGPFTDVVTADLKLSNPTDRRICFKVKTTAPKRYCVRPNSGI
HP IDPGSTVTVSVMLQPFDYDPNEKSKHKFMVQTIFAPPNTSD-MEAVWKEAKPDELMDSKL
   ..* ....*.******.******.*******...** .. .  * .**.* *..***.**
AP LEPKTSIAVAVMLQPFNYDPNEKNKHKFMVQSMYAPDHVVESQELLWKDAPPESLMDTKL
HP RCVFEMPNENDKLNDMEPSK---------------AVPLNASKQDGPMPKP-HSVSLNDTE
   *******.....  . ..*.                . .....* ... **.  .*. ....
AP RCVFEMPDGSHQAPASDASRATDAGAHFSESALEDPTVASRKTETQSPKRVGAVGSAGED
HP TRKLMEECKRLQGEMMKLSEENRHLRDEGLRLRKVAHSD--KPGSTSTASFRDNVTSPLP
   ..** .* *. *.*. .*..**..*.***.****** .*  .*.. .... ........*
AP VKKLQHELKKAQSEITSLKGENSQLKDEGIRLRKVAMTDTVSPTPLNPSPAPAAAVRAFP
HP SLLVVIAAIFIGFFLGKFIL
   ... *.***..*...***.*
AP PVVYVVAAIILGLIIGKFLL
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA447905) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the

<HP10423> (SEQ ID Nos. 4, 14, and 24)

Determination of the whole base sequence of the cDNA insert of clone HP10423 obtained from cDNA library of human osteosarcoma cell line U-2 OS revealed the structure consisting of a 64-bp 5'-untranslated region, a 795-bp ORF, and a 207-bp 3'-untranslated region. The ORF codes for a protein consisting of 264 amino acid residues and there existed a secretory signal at the N-terminus and one putative transmembrane domain at the N-terminus. Figure 4 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 30 kDa that was almost identical with the molecular weight of 29,377 predicted from the ORF. When expressed in COS7 cells, an expression product of about 31 kDa was observed in the membrane fraction.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. D80116) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10506> (SEQ ID Nos. 5, 15, and 25)

Determination of the whole base sequence of the cDNA insert of clone HP10506 obtained from cDNA library of human stomach cancer revealed the structure consisting of a 53-bp 5'-untranslated region, a 339-bp ORF, and a 226-bp 3'-untranslated region. The ORF codes for a protein consisting [illegible] transmembrane domain. Figure [illegible] depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-

Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 12 kDa that was almost identical with the molecular weight of 11,821 predicted from the ORF. When expressed in COS7 cells, an expression product of about 13 kDa was observed in the membrane fraction.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA282544) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10507> (SEQ ID Nos. 6, 16, and 26)

Determination of the whole base sequence of the cDNA insert of clone HP10507 obtained from cDNA library of human stomach cancer revealed the structure consisting of a 412-bp 5'-untranslated region, a 441-bp ORF, and a 168-bp 3'-untranslated region. The ORF codes for a protein consisting of 146 amino acid residues and there existed a secretory signal at the N-terminus and one putative transmembrane domain at the C-terminus. Figure 6 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 19 kDa that was somewhat larger than the molecular weight of 16,347 predicted from the ORF.

Furthermore, the search of the GenBank using the base [illegible] of sequences that shared a homology of 90% or more (for example, Accession No. AA424759) in ESTs, but, since they

are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10548> (SEQ ID Nos. 7, 17, and 27)

Determination of the whole base sequence of the cDNA insert of clone HP10548 obtained from cDNA library of human stomach cancer revealed the structure consisting of a 330-bp 5′-untranslated region, a 1035-bp ORF, and a 67-bp 3′-untranslated region. The ORF codes for a protein consisting of 344 amino acid residues and there existed four putative transmembrane domains. Figure 7 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of a high molecular weight.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA143152) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10566> (SEQ ID Nos. 8, 18, and 28)

Determination of the whole base sequence of the cDNA insert of clone HP10566 obtained from cDNA library of the human stomach cancer revealed the structure consisting of a 61-bp 5′-untranslated region, a 294-bp ORF, and a 246-bp 3′- [illegible] [illegible] amino acid residues and there existed one putative transmembrane domain at the C-terminus. Figure 8 depicts the

hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 12 kDa that was almost identical with the molecular weight of 11,452 predicted from the ORF. When expressed in COS7 cells, an expression product of about 12 kDa was observed in the membrane fraction.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. W79821) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10567> (SEQ ID Nos. 9, 19, and 29)

Determination of the whole base sequence of the cDNA insert of clone HP10567 obtained from cDNA library of the human stomach cancer revealed the structure consisting of a 77-bp 5'-untranslated region, a 375-bp ORF, and a 133-bp 3'-untranslated region. The ORF codes for a protein consisting of 124 amino acid residues and there existed one putative transmembrane domain at the C-terminus. Figure 9 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 14 kDa that was almost identical with the molecular weight of 14,484 predicted from the ORF.

Furthermore, the search of the GenBank using the base [illegible] of sequences that shared a homology of 90% or more (for example, Accession No. AA428475) in ESTs, but, since they

are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10568> (SEQ ID Nos. 10, 20, and 30)

Determination of the whole base sequence of the cDNA insert of clone HP10568 obtained from cDNA library of the human stomach cancer revealed the structure consisting of a 56-bp 5'-untranslated region, a 984-bp ORF, and a 60-bp 3'-untranslated region. The ORF codes for a protein consisting of 327 amino acid residues and there existed a secretory signal at the N-terminus and one putative transmembrane domain at the C-terminus. Figure 10 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 36.5 kDa that was almost identical with the molecular weight of 34,326 predicted from the ORF. In this case, the addition of a microsome led to the formation of a product of 40 kDa which is considered to have a sugar chain being attached. In addition, there exist in the amino acid sequence of this protein two sites at which N-glycosylation may occur (Asn-Leu-Thr at position 138 and Asn-Leu-Ser at position 206). Application of the (-3,-1) rule, a method for predicting the cleavage site of the secretory signal sequence, allows to expect that the mature protein starts from valine at position 24. When expressed in COS7 cells, an expression product of about 31 kDa was observed in the supernatant fraction and the membrane fraction.

acid sequence of the present protein has revealed that the protein was similar to the human cell-surface A33 antigen

(SWISS-PROT Accession No. Q99795). Table 5 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the human cell-surface A33 antigen (A3). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 30.0% in the N-terminal region of 243 residues.

Table 5

```
HP MAELPGPFLCGALLGFLCLSGLAVEVKVPTEPLSTPLGKTAELTCTYSTSVGDSFAL-EW
                                  *..*..* . *... **...*.*** **.... .* .*
A3   MVGKMWPVLWTLCAVRVTVDAISVETPQDVLRASQGKSVTLPCTYHTSTSSREGLIQW
HP SFVQPGKPISESHPILYFTNGHLYPTGSKSKRVSLLQNPPTVGVATLKLTDVHPSDTGTY
   . .     . .*.  *  *.* . *  *.  *.    . * .. . *.......  .*.***
A3 DKLL--LTHTERVVIWPFSNKN-YIHGELYKNRVSISNNAEQSDASITIDQLTMADNGTY
HP LCQVNNPPDFYTNGLGLINLTVLVPPSNPLCSQSGQTSVGGSTALRCSSSEGAPKPVYNW
    * *.   .*. .*. . ..* ******.* *. .*.* .*....* * *.**.*.. ..*
A3 ECSVSLMSDLEGNTKSRVRLLVLVPPSKPECGIEGETIIGNNIQLTCQSKEGSPTPQYSW
HP VRLGTFPTPSPGSMVQDEVSGQLILTNLSLTSSGTYRCVATHQMGSASCELTLSVTEPS-
    * ...    *  ..* . . .. *.*.* ..** * *...*. *.. *..*..* .**
A3 KRYNILNQEQP--LAQPASGQPVSLKNISTDTSGYYICTSSNEEGTQ[illegible]NITVAV[illegible]SM
HP -QGRVAGALIGVLLGVLLLSVAAFCLVRFQKERGKKPKEFYGGSDLREDAIAPG[illegible]TC
     .  .*  .**. .......  .*
A3 NVALYVGIAVGVVAALIIIGIIIYCCCCRGKDDNTEDKEDARPNREAYEEPPEQ[illegible]
HP MRADSSKGFLERPSSASTVTTTKSKLPMVV

A3 EREEEDDYRQEEQRSTGRESPDHLDQ
```

Furthermore, the search [illegible] the [illegible] sequences of the present cDNA has revealed [illegible]

of sequences that shared a homology of 90% or more (for example, Accession No. T24595) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP01426> (SEQ ID Nos. 31, 41, and 51)

Determination of the whole base sequence of the cDNA insert of clone HP01426 obtained from cDNA library of human stomach cancer revealed the structure consisting of a 1-bp 5'-untranslated region, a 942-bp ORF, and a 122-bp 3'-untranslated region. The ORF codes for a protein consisting of 313 amino acid residues and there existed a putative secretory signal. Figure 11 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 36 kDa that was almost identical with the molecular weight of 34,955 predicted from the ORF. In this case, the addition of a microsome led to the formation of a product of 38 kDa which is considered to have a sugar chain being attached after secretion. In addition, there exists in the amino acid sequence of this protein one site at which N-glycosylation may occur (Asn-Ser-Ser at position 163). Application of the (-3,-1) rule, a method for predicting the cleavage site of the secretory signal sequence, allows to expect that the mature protein starts from tryptophan at position 17. When expressed in COS7 cells, an expression product of about 39 kDa was observed in the supernatant [illegible]

The search of the protein data base using the amino acid sequence of the present protein revealed that the

protein was similar to the Xenopus laevis cortical granule lectin (EMBL Accession No. X82626). Table 6 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the X. laevis cortical granule lectin (XL). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 67.9% in the region other than the N-terminal region.

Table 6

```
HP MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRT
   *     **                  *           ********.  .* **.* * .
XL MLVHILLLLVTGGLSQSCEPVVIVASKNMVKQLDCDKFRSCKEIKDSNEEAQDGIYTLTS
HP ENGVIYQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANY
   ..*. ********..*************.* ****.*********..*************
XL SDGISYQTFCDMTTNGGGWTLVASVHENNMAGKCTIGDRWSSQQGNRADYPEGDGNWANY
HP NTFGSAEAATSDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLG
   ******..**************.* .**.******.*.  ****** ****..*.* . *
XL NTFGSAGGATSDDYKNPGYYDIEAYNLGVWHVPNKTPLSVWRNSSLQRYRTTDGILFKHG
HP HNLFGIYQKYPVKYGEGKCWTDNGPVIPVVYDFGDAQKTASYYSPYGQREFTAGFVQFRV
    ***..*. ****** *.* .*.**..*****.*.*. ***.***  ...**.*..***
XL GNLFSLYRIYPVKYGIGSCSKDSGPTVPVVYDLGSAKLTASFYSPDFRSQFTPGYIQFRP
HP FNNERAANALCAGMRVTGCNTEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSS
   .*.*.** ***.**....**.** ***********.*.*****...*..****.   *..
XL INTEKAALALCPGMKMESCNVEHVCIGGGGYFPEADPRQCGDFAAYDFNGYGTKKFNSAG
HP REITEAAVLLFYR
   ************
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. R06009) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP02515> (SEQ ID Nos. 32, 42, and 52)

Determination of the whole base sequence of the cDNA insert of clone HP02515 obtained from cDNA library of human osteosarcoma cell line Saos-2 revealed the structure consisting of a 176-bp 5'-untranslated region, a 690-bp ORF, and a 71-bp 3'-untranslated region. The ORF codes for a protein consisting of 229 amino acid residues and there existed a putative secretory signal at N-terminus and one putative transmembrane domain at the C-terminus. Figure 12 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 27 kDa that was almost identical with the molecular weight of 26,000 predicted from the ORF. In this case, the addition of a microsome led to the formation of a product of 25.5 kDa from which the secretory signal is considered to have been cleaved. Application of the (-3,-1) rule, a method for predicting the cleavage site of the secretory signal sequence, allows to expect that the mature protein starts from phenylalanine at position 28.

The search of the protein data base using the amino [illegible] protein was similar to the human T1/ST2 receptor binding protein (GenBank Accession No. U41804). Table 7 shows the

comparison between amino acid sequences of the human protein of the present invention (HP) and the human T1/ST2 receptor binding protein (T1). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 55.8% in the entire region.

Table 7

```
HP  MGDKIWLPFPVLLLAALPPVLLPGAAGFTPSLDSDFTFTLPAGQKECFYQPMPLKASLE
          *.... **  .*** . *.** .*  *..*** ****.*.****. * .****
T1 MMAAGAALALALWLL--MPPVEV-GGAGPPPIQDGEFTFLLPAGRKQCFYQSAPANASLE
HP IEYQVLDGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVE-TEVGDYMFCFDNTFSTISEK
   .****..*****.** *.**.*  ** * **.******* **.***..****.*******
T1 TEYQVIGGAGLDVDFTLESPQGVLLVSESRKADGVHTVEPTEAGDYKLCFDNSFSTISEK
HP VIFFELILDNMGEQAQEQEDWKKYITGTDILDMKLEDILESINSIKSRLSKSGHIQILLR
   ..*****.*.. ....* *.* . .. ...**.*.*** ***.....**..* .. .***
T1 LVFFELIFDSL-QDDEEVEGWAEAVEPEEMLDVKMEDIKESIETMRTRLERSIQMLTLLR
HP AFEARDRNIQESNFDRVNFWSMVNLVVMVVVSAIQVYMLKSLFEDKRKSRT
   ********.**.*..****** **..*...*...**. **..*.***  .*
T1 AFEARDRNLQEGNLERVNFWSAVNVAVLLLVAVLQVCTLKRFFQDKRPVPT
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA381943) in ESTs, but, since they

[illegible]

protein of the present invention.

<HP02575> (SEQ ID Nos. 33, 43, and 53)

Determination of the whole base sequence of the cDNA insert of clone HP02575 obtained from cDNA library of human osteosarcome cell line Saos-2 revealed the structure consisting of a 55-bp 5'-untranslated region, a 1404-bp ORF, and a 219-bp 3'-untranslated region. The ORF codes for a protein consisting of 467 amino acid residues and there existed a putative secretory signal at the N-terminus. Figure 13 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 52 kDa that was almost identical with the molecular weight of 54,065 predicted from the ORF. In this case, the addition of a microsome led to the formation of a product of 57 kDa which is considered to have a sugar chain being attached afetr secretion. In addition, there exist in the amino acid sequence of this protein three sites at which N-glycosylation may occur (Asn-Arg-Thr at position 171, Asn-Ser-Thr at position 239 and Asn-Asp-Thr at position 377). Application of the (-3,-1) rule, a method for predicting the cleavage site of the secretory signal sequence, allows to expect that the mature protein starts from histidine at position 29. When expressed in COS7 cells, an expression product of about 55 kDa was observed in the supernatant fraction.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the human α-L-fucosidase (SWISS-PROT [illegible] amino acid sequences of the human protein of the present invention (HP) and the human α-L-fucosidase (FC). Therein,

the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 54.8% in the entire region.

Table 8

```
HP MRPQELPRLAFPLLLLLLLLLPPPPC-PAHSATRFDPTWESLDARQLPAWFDQAKFGIFI
                     .******.* .. .  .*... *..*.* ***.*.******.****.**
FC    MRSRPAGPALLLLLLFLGAAESVRRAQPPRRYTPDWPSLDSRPLPAWFDEAKFGVFI
HP HWGVFSVPSFGSEWFWWYWQKEKIPKYVEFMKDNYPPSFKYEDFGPLFTAKFFNANQWAD
   ********..*******.** *   *.*  **.*****.*.*.**** ***.**....***
FC HWGVFSVPAWGSEWFWWHWQGEGRPQYQRFMRDNYPPGFSYADFGPQFTARFFHPEEWAD
HP IFQASGAKYIVLTSKHHEGFTLWGSEYSWNWNAIDEGPKRDIVKELEVAIRNRTDLRFGL
   .***.****.***.******* * *  *****. * **.**.* **..*.*.* ..*.**
FC LFQAAGAKYVVLTTKHHEGFTNWPSPVSWNWNSKDVGPHRDLVGELGTALRKR-NIRYGL
HP YYSLFEWFHPLFLEDESSSFHKRQFPVSKTLPELYELVNNYQPEVLWSDGDGGAPDQYWN
   *.**.******.* *....*....*  .**.****.***.*.*...****. . ** ***
FC YHSLLEWFHPLYLLDKKNGFKTQHFVSAKTMPELYDLVNSYKPDLIWSDGEWECPDTYWN
HP STGFLAWLYNESPVRGTVVTNDRWGAGSICKHGGFYTCSDRYNPGHLLPHKWENCMTIDK
   **.**.****.***...**.*****... *.***.*.*.*...*  *  **** * .***
FC STNFLSWLYNDSPVKDEVVVNDRWGQNCSCHHGGYYNCEDKFKPQSLPDHKWEMCTSIDK
HP LSWGYRREAGISDYLTIEELVKQLVETVSCGGNLLMNIGPTLDGTISVVFEERLRQMGSW
   .******. ..**  . .*....**.*** *** *.***** ** *  .*.*** ..*.*
FC FSWGYRRDMALSDVTEESEIISELVQTVSLGGNYLLNIGPTKDGLIVPIFQERLLAVGKW
HP LKVNGEAIYETHTWRSQNDTVTPDVWYTSKPKEKLVYAIFLKWPTSGQLFLGHPKAILGA
   *..******....** * .. *..****** ..  ******.**..* * *. * .. ..
FC LSINGEAIYASKPWRVQWEKNTTSVWYTSKGSA--VYAIFLHWPENGVLNLESPITT-ST
HP TEVKLLGHGQPLNWISLEQNGIMVELPQLTIHQMPCKWGWALALTNVI
   .  ..   . .    .    ....    .    .   ..
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. N28668) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10357> (SEQ ID Nos. 34, 44, and 54)

Determination of the whole base sequence of the cDNA insert of clone HP10357 obtained from cDNA library of human stomach cancer revealed the structure consisting of a 113-bp 5′-untranslated region, a 300-bp ORF, and a 54-bp 3′-untranslated region. The ORF codes for a protein consisting of 99 amino acid residues and there existed two putative transmembrare domains. Figure 14 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 11 kDa that was almost identical with the molecular weight of 10,923 predicted from the ORF.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA477156) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10447> (SEQ ID Nos. 35, 45, and 55)

Determination of the whole base sequence of the cDNA

insert of clone HP10447 obtained from cDNA library of human liver revealed the structure consisting of a 271-bp 5'-untranslated region, a 570-bp ORF, and a 34-bp 3'-untranslated region. The ORF codes for a protein consisting of 189 amino acid residues and there existed five putative transmembrare domains. Figure 15 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of high molecular weight.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA296976) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10477> (SEQ ID Nos. 36, 46, and 56)

Determination of the whole base sequence of the cDNA insert of clone HP10477 obtained from cDNA library of human liver revealed the structure consisting of a 149-bp 5'-untranslated region, a 1092-bp ORF, and a 15-bp 3'-untranslated region. The ORF codes for a protein consisting of 363 amino acid residues and there existed one putative transmembrane domain at the N-terminus. Figure 16 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product [illegible] weight of 39,884 predicted from the ORF.

The search of the protein data base using the amino

acid sequence of the present protein revealed that the protein was similar to the human peptidoglycan recognition protein (GenBank Accession No. AF076483). Table 9 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the human peptidoglycan recognition protein (PG). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 54.8% in the entire region.

Table 9

```
HP MVDSLLAVTLAGNLGLTFLRGSQTQSHPDLGTEGCWDQLSAPRTFTLLDPKASLLTKAFL
HP NGALDGVILGDYLSRTPEPRPSLSHLLSQYYGAGVARDPGFRSNFRRQNGAALTSASILA
HP QQVWGTLVLLQRLEPVHLQLQCMSQEQLAQVAANATKEFTEAFLGCPAIHPRCRWGAAPY
                                                *..* **  * * .
PG                 MSRRSMLLAWALPSLLRLGAAQETEDPACCSPIVPRNEWKALA-
HP RGRPKLLQLPLGFLYVHHTYVPAPPCTDFTRCAANMRSMQRYHQDTQGWGDIGYSFVVGS
   .. .. * *** .. * **  ....*.. ..*... *..*.** .* ** *.**.*..*.
PG SECAQHLSLPLRYVVVSHT--AGSSCNTPASCQQQARNVQHYHMKTLGWCDVGYNFLIGE
HP DGYVVYEGRGWHWVGAHTLGH-NSRGFGVAIVGNYTAALPTEAALRTVRDTLPSCAVRAGL
   ** *******...***.    *....*....*** . .** .*.*.... * .*.* .*
PG DGLVYEGRGWNFTGAHSGHLWNPMSIGISFMGNYMDRVPTPQAIRAAQGLL-ACGVAQGA
HP LRPDYALLGHRQLVRTDCPGDALFDLLRTWPHFTATVKPRPARSVSKRSRREPPPRTLPA
   **..*.* ***.. ** .**..*..*...***.
PG LRSNYVLKGHRDVQRTLSPGNQLYHLIQNWPHYRSP
```

sequences of the present cDNA has revealed the registration

of sequences that shared a homology of 90% or more (for example, Accession No. AA424759) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10513> (SEQ ID Nos. 37, 47, and 57)

Determination of the whole base sequence of the cDNA insert of clone HP10513 obtained from cDNA library of human stomach cancer revealed the structure consisting of a 134-bp 5'-untranslated region, a 750-bp ORF, and a 0-bp 3'-untranslated region. The ORF codes for a protein consisting of 249 amino acid residues and there existed one putative transmembrane domain at the N-terminus. Figure 17 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 29 kDa that was almost identical with the molecular weight of 27,373 predicted from the ORF.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the human hypothetical protein KIAA0512 (GenBank Accession No. AB011084). Table 10 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the human hypothetical protein KIAA0512 (KI). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the [illegible] homology of 31.6% in the C-terminal region of 196 amino acid residues.

Table 10

```
HP                           MGGPRGAGWVAAGLLLGAGACYCIYRLTRGRRRG

KI RGRGRRPVAMQKRPFPYEIDEILGVRDLRKVLALLQKSDDPFIQQVALLTLSNNANYSCN
HP DRELGIRSSKSAEDLTDGSYDDVLNAEQLQKLLYLLESTEDPVIIERALITLGNNAAFSV
                       *  .....*  .  * *. *..  ..  . .. ..   *
KI QETIRKLGGLPIIANMINKTDPHIKEKALMAMNNLSENYENQGRLQVYMNKVMDDIMASN
HP NQAIIRELGGIPIVANKINHSNQSIKEKALNALNNLSVNVENQIKIKVQVLKLLLNLSEN
    .. .. .*   ... .*... * .  ......  ...  ..  ****..**.* *..**
KI LNSAVQVVGLKFLTNMTITNDYQHLLVNSIANF--FRLLSQGGGKIKVEILKILSNFAEN
HP PAMTEGLLRAQVDSSFLSLYDSHVAKEILLRVLTLFQNIKNCLKIEGHLAVQPTFTEGSL
   *.* . **..** .** ***.*.*..***...****. * . *. *  .   ..*..***
KI PDMLKKLLSTQVPASFSSLYNSYVESEILINALTLFEIIYDNLRAE--VFNYREFNKGSL
HP FFL-LHGEECAQKIRALVDHHDAEVKEKVVTIIPKI
   *.*   .. *..*****..***  ** **.... *.
KI FYLCTTSGVCVKKIRALANHHDLLVKVKVIKLVNKF
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. N92228) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10540> (SEQ ID Nos. 38, 48, and 58)

osteosarcoma cell line Saos-2 revealed the structure

consisting of a 47-bp 5'-untranslated region, a 297-bp ORF, and a 245-bp 3'-untranslated region. The ORF codes for a protein consisting of 98 amino acid residues and there existed two putative transmembrane domains. Figure 18 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of high molecular weight.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Caenorhabditis elegans hypothetical protein CEF49C12.12 (GenBank Accession No. Z68227). Table 11 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the C. elegans hypothetical protein CEF49C12.12 (CE). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 36.1% in the entire region.

Table 11

```
HP M-ASLLCCGPKLAACGIVLSAWGVIMLIMLGIFFNVHSAVLIEDVPFTEKDFENGPQNIY
   *       ***  *   * * **** *  ** **      *  *  * *         *
CE MGKICPLMGPKMSAFCMVMSVWGVIFLGLLGVFFYIQAVTLFPDLHF-EGHGKVPSSVID
HP NLYEQVSYNCFIAAGLYLLLGGFSFCQVRLNKRKEYMVR
     *      * ******       * *   **
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA420715) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10557> (SEQ ID Nos. 39, 49, and 59)

Determination of the whole base sequence of the cDNA insert of clone HP10557 obtained from cDNA library of human stomach cancer revealed the structure consisting of a 24-bp 5′-untranslated region, a 519-bp ORF, and a 130-bp 3′-untranslated region. The ORF codes for a protein consisting of 172 amino acid residues and there existed a putative secretory signal at the N-terminus. Figure 19 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 32 kDa that was larger than the molecular weight of 18,844 predicted from the ORF. In this case, the addition of a microsome led to the formation of a product of 39 kDa which is considered to have been subjected to some modification after secretion. In addition, there exist in the amino acid sequence of this protein no site at which N-glycosylation may occur. Application of the (-3,-1) rule, a method for predicting the cleavage site of the secretory signal sequence, allows to expect that the mature protein [illegible] cells, an expression product of about [illegible] kDa was observed in the supernatant fraction and the membrane fraction.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the human progesterone binding protein (EMBL Accession No. AJ002030). Table 12 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the human progesterone binding protein (PG). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 30.5% in the C-terminal region of 151 amino acid residues.

Table 12

```
HP                                               MVGPAP

PG MAAGDGDVKLGTLGSGSESSNDGGSESPGDAGAAAEGGGWAAAALALLTGGGEMLLNVAL
HP RRRLRPLAALALVLALAPGLPTARAGQTPRPAERGPPV--RLFTEEELARYGGEEEDQPI
                  ** .. .. . .**.. *..  * *. *.* .*.* . ...*
PG VALVLLGAYRLWVRWGRRGLGAGAGAGEESPATSLPRMKKRDFSLEQLRQYDG-SRNPRI
HP YLAVKGVVFDVTSGKEFYGRGAPYNALTGKDSTRGVAKMSLDPADLTHDTTGLTAKELEA
     ***.* *****.*..***...**. ..*.*..**.*...** ..* .. ..*.. .  .
PG LLAVNGKVFDVTKGSKFYGPAGPYGIFAGRDASRGLATFCLDKDALRDEYDDLSDLNAVQ
HP LDEV--FTKVYKAKYPIVGYTARRILNEDGSPNLDFKPEDQPHFDIKDEF
   ...*  ... .*.**  .*.. *.*. ...*. ... *.... . *..
PG MESVREWEMQFKEKY---DYVG-RLLKPGEEPS-EYTDEEDTKDHNKQD
```

of sequences that shared a homology of 90% or more (for

example, Accession No. AA101709) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10563> (SEQ ID Nos. 40, 50, and 60)

Determination of the whole base sequence of the cDNA insert of clone HP10563 obtained from cDNA library of human osteosarcoma cell line Saos-2 revealed the structure consisting of a 126-bp 5'-untranslated region, a 363-bp ORF, and a 936-bp 3'-untranslated region. The ORF codes for a protein consisting of 120 amino acid residues and there existed two putative transmembrane domains. Figure 20 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 18.5 kDa that was larger than the molecular weight of 13,180 predicted from the ORF.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Arabidopsis thaliana hypothetical protein F27F23.15 (GenBank Accession No. AC003058). Table 13 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the A. thaliana hypothetical protein F27F23.15 (AT). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins

Table 13

```
HP MMPSRTNLATGIPSSKVKYSRLSSTDDGYIDLQFKKTPPKIPYKAIALATVLFLIGAFLI
                                        *..*  *.  . ... . * *.*.**. *...*.. *
AT             MAYVDHAFSISDEDLMIGTSY-TVSNRPPVKEISLAVGLLVFGTLGI
HP IIGSLLLSGYISKGGADRAVPVLIIGILVFLPGFYHLRIAYYASKGYRGYSYDDIPDFDD
   ..* ..  . .. *. ....  ...* *.*.****. ****** ***.*.*...**
AT VLGFFMAYNRVG-GDRGHGIFFIVLGCLLFIPGFYYTRIAYYAYKGYKGFSFSNIPSV
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA083574) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP01467> (SEQ ID Nos. 61, 71, and 81)

Determination of the whole base sequence of the cDNA insert of clone HP01467 obtained from cDNA library of human fibrosarcoma cell line HT-1080 revealed the structure consisting of a 65-bp 5′-untranslated region, a 924-bp ORF, and a 447-bp 3′-untranslated region. The ORF codes for a protein consisting of 307 amino acid residues and there existed three putative transmembrane domains. Figure 21 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation

The search of the protein data base using the amino

acid sequence of the present protein revealed that the protein was similar to the rat Sec22 homologue (GenBank Accession No. U42209). Table 14 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the rat Sec22 homologue (RN). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 94.6% in the N-terminal region of 241 amino acid residues. The protein of the present invention was longer by 53 amino acids at the C-terminus than the rat Sec22 homologue.

Table 14

```
HP MSMILSASVIRVRDGLPLSASTDYEQSTGMQECRKYFKMLSRKLAQLPDRCTLKTGHYNI
   *********.*************.***.*.****************.*********..**
RN MSMILSASVVRVRDGLPLSASTDCEQSAGVQECRKYFKMLSRKLAQFPDRCTLKTGRHNI
HP NFISSLGVSYMMLCTENYPNVLAFSFLDELQKEFITTYNMMKTNTAVRPYCFIEFDNFIQ
   ************************************************************
RN NFISSLGVSYMMLCTENYPNVLAFSFLDELQKEFITTYNMMKTNTAVRPYCFIEFDNFIQ
HP RTKQRYNNPRSLSTKINLSDMQTEIKLRPPYQISMCELGSANGVTSAFSVDCKGAGKISS
   ********************** **********.**************************
RN RTKQRYNNPRSLSTKINLSDMQMEIKLRPPYQIPMCELGSANGVTSAFSVDCKGAGKISS
HP AHQRLEPATLSGIVGFILSLLCGALNLIRGFHAIESLLQSDGDDFNYIIAFFLGTAACLY
   **************.***************************.**.*.************
RN AHQRLEPATLSGIVAFILSLLCGALNLIRGFHAIESLLQSDGEDFSYMIAFFLGTAACLY
HP QCYLLVYYTGWRNVKSFLTFGLICLCNMYLYELRNLWQLFFHVTVGAFVTLQIWLRQAQG
   *
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA421925) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP01956> (SEQ ID Nos. 62, 72, and 82)

Determination of the whole base sequence of the cDNA insert of clone HP01956 obtained from cDNA library of human liver revealed the structure consisting of a 86-bp 5'-untranslated region, a 552-bp ORF, and a 359-bp 3'-untranslated region. The ORF codes for a protein consisting of 183 amino acid residues and there existed one putative transmembrane domain. Figure 22 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 20.5 kDa that was almost identical with the molecular weight of 20,073 predicted from the ORF.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the yeast hypothetical protein 21.5 kDa (SWISS-PROT Accession No. P53073). Table 15 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the yeast hypothetical protein 21.5 kDa (SC). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that

[illegible]

invention, respectively. The both proteins shared [illegible]

of 34.3% in the C-terminal region of 108 amino acid residues.

Table 15

```
HP      MTAQGGLVANRGRRFKWAIELSGPGGGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS

SC MSEQEPYEWAKHLLDTKYIEKYNIQNSNTLPSPPGFEGNSSKGNVTRKQQDATSQTTSLA
HP VQETDRILVEKRCWDIALGPLKQIPMNLFIMYMAGNTISIFPTMMVCMMAWRPIQALMAI
          .*  .. *.*** * * ****.*. **.*... *.*.* . *.   **.*....
SC QKNQITVLQVQKAWQIALQPAKSIPMNIFMSYMSGTSLQIIPIMTALMLLSGPIKAIFST
HP SATFK--MLESSSQKFLQGLVYLIGNLMCLALAV-Y-KCQSMGLLPTHASDWLAFIEPPE
   ...**  . ....*. .*. ...   . *. . . * * .****.*.  .***.
SC RSAFKPVLGNKATQSQVQTAMFMYIVFQGVLMYIGYRKLNSMGLIPNAKGDWLPWERIAH
HP RMEFSGGGLLL

SC YNNGLQWFSD
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA159753) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP02545> (SEQ ID Nos. 63, 73, and 83)

Determination of the whole base sequence of the cDNA insert of clone HP02545 obtained from cDNA library of human [illegible] and a 636-bp [illegible]-untranslated region. The ORF codes for a

protein consisting of 327 amino acid residues and there existed a putative secretory signal at the N-terminus and one putative transmembrane domain at the C-terminus. Figure 23 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the rat embigin (EMBL Accession No. AJ009698). Table 16 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the rat embigin (RN). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 65.4% in the entire region.

Table 16

```
HP MRALPGLLEARARTPRLLLLQCLLAAARPSSADGSAPDSPFTSPPLREEIMAN--NFSLE
   **. .** .  *.. .**** .****.**..* *...**.*** *.***.**.  *.***
RN MRSHTGLRALVAPGCSLLLL-YLLAATRPDRAVGDPADSAFTSLPVREEMMAKYANLSLE
HP SHNISLTEHSSMPVEKNITLERPSNVNLTCQFTTSGDLNAVNVTWKKDGEQLE--NNYLV
   ..******..... *.********...*.* **...*. ..*******.. **  ... .
RN TYNISLTEQTRVS-EQNITLERPSHLELECTFTATEDVMSMNVTWKKDDALLETTDGFNT
HP SATGSTLYTQYRFTIINSKQMGSYSCFFREEKEQRGTFNFKVPELHGKNKPLISYVGDST
   .  *.***.*****..******.****. **   *****..**..********.******
RN TKMGDTLYSQYRFTVFNSKQMGKYSCFLGEE--LRGTFNIRVPKVHGKNKPLITYVGDST
HP VLTCKCQNCFPLNWTWYSSNGSVKVPVGVQM-NKYVINGTYANETKLKITQLLEEDGESY
   **.*.****.******* ***...**..*.. .*. ***.********...******.**
RN VLKCECQNCLPLNWTWYMSNGTAQVPIDVHVNDKFDINGSYANETKLKVKHLLEEDGGSY
HP WCRALFQLGESEEHIELVVLSYLVPLKPFLVIVAEVILLVATILLCEKYTQKKKKHSDEG
   **** *.********.*****..*******.*.********.***** ******...*.*
RN WCRAAFPLGESEEHIKLVVLSFMVPLKPFLAIIAEVILLVAIILLCEVYTQKKKNDPDDG
HP KEFEQIEQLKSDDSNGIENNVPRHRKNESLGQ
   ***********************.**..* .*
RN KEFEQIEQLKSDDSNGIENNVPRYRKTDSGDQ
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA312629) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

Determination of the whole base sequence of the cDNA insert of clone HP02551 obtained from cDNA library of human

osteosarcoma cell line Saos-2 revealed the structure consisting of a 61-bp 5′-untranslated region, a 672-bp ORF, and a 384-bp 3′-untranslated region. The ORF codes for a protein consisting of 223 amino acid residues and there existed a putative secretory signal at the N-terminus. Figure 24 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 27 kDa that was somewhat larger than the molecular weight of 24,555 predicted from the ORF. In this case, the addition of a microsome led to the formation of a product of 26 kDa from which the secretory signal is considered to have been cleaved. Application of the (-3,-1) rule, a method for predicting the cleavage site of the secretory signal sequence, allows to expect that the mature protein starts from glutamine at position 20.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the mouse FGF binding protein (GenBank Accession No. U49641). Table 17 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the mouse FGF binding protein (MM). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 21.2% in the entire region other than the N-terminal region. In particular, all the eight cysteine residues contained in the

Table 17

```
HP                                  MKFVPCLLLVTLSCLGTLGQAPRQKQGST
                                                ..**.  . .* . ...
MM MRLHSLILLSFLLLATQAFSEKVRKRAKNAPHSTAEEGVEGSAPSLGKAQNKQRSRTSKS
HP GEEFHFQTGGRDSCTMRPSSLGQGAGEVWLRVDCRNTDQTYWCEYRGQPSMCQAFAADPK
    .. .* * ....*   . ...... .. *.*.* ..**.. * . *.*. *   . * .
MM LTHGKFVTKDQATC---RWAVTEEEQGISLKVQCTQADQEFSCVFAGDPTDCLKHDKD-Q
HP SYWNQALQELRRLHHACQGA-PVLRPSVCREAGPQAHMQQVTSSLKGSPEPNQQPEAGTP
   **.*. ..**. .. *..* .**...***.  *..... *... .*...*... .. ..
MM IYWKQVARTLRKQKNICRDAKSVLKTRVCRKRFPESNLKLVNPNARGNTKPRKEKAEVSA
HP SLRPKATVKLTEATQLGKDSMEELGKAKPTTRPTAKPTQPGPRPGGNEEAKKKAWEHCWK
   . . *.   ......  *... .*. * . *. * . ..    *.  .. .*.* * * .
MM REHNKVQEAVSTEPNRIKEDI-TLNPAATQTM-TIRDPECLEDPDVLNQ-RKTALEFCGE
HP PFQALCAFLISFFRG
   .. ..*.*.......
MM SWSSICTFFLNMLQATSC
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA317400) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP02631> (SEQ ID Nos. 65, 75, and 85)

Determination of the whole base sequence of the cDNA insert of clone HP02631 obtained from cDNA library of human [illegible] consisting of a 42-bp 5'-untranslated region, a 147-bp ORF,

and a 1191-bp 3'-untranslated region. The ORF codes for a protein consisting of 48 amino acid residues and there existed a putative secretory signal at the N-terminus. Figure 25 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 10 kDa or less.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA156969) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP02632> (SEQ ID Nos. 66, 76, and 86)

Determination of the whole base sequence of the cDNA insert of clone HP02632 obtained from cDNA library of human fibrosarcoma cell line HT-1080 revealed the structure consisting of a 50-bp 5'-untranslated region, a 1116-bp ORF, and a 337-bp 3'-untranslated region. The ORF codes for a protein consisting of 371 amino acid residues and there existed eight putative transmembrane domains. Figure 26 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of high molecular weight.

The search of the protein data base using the amino acid sequence of the present protein revealed that the

hypothetical protein CELC2H... GenBank Accession N...

Table 18 shows the comparison between amino acid ...

of the human protein of the present invention (HP) and the C. elegans hypothetical protein CELC2H12 (CE). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 51.4% in the entire region.

Table 18

```
HP MAWTKYQLFLAGLMLVTGSINTLSAKWADNFMAEGCGGSKEHSFQHPFLQAVGMFLGEFS
        ...... .*.****.**..*****...*.      . .*.******. **.***
CE   MVAFAVIISVMMVVTGSLNTICAKWADSIKAD------GVPFNHPFLQATCMFFGEFL
HP CLAAFYL---------LRCRAAGQSDS-------SVDPQQPFNPLLFLPPALCDMTGTSL
   **..*.*          * ...*.*.*       .  . .****.**.******. ***.
CE CLVVFFLIFGYKRYVWNRANVQGESGSVTEITSEEKPTLPPFNPFLFFPPALCDILGTSI
HP MYVALNMTSASSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWLGILATIAGLVVVGLADLL
   **..**.*.*****************.**..*. ..   .*.*.* .. ***.**..*.
CE MYIGLNLTTASSFQMLRGAVIIFTGLLSVGMLNAQIKPFKWFGMLFVMLGLVIVGVTDIY
HP SKHDSQHKLSEVITGDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLRAVGTEGLFGFVILS
    ..*. .. ...***.***.*********** *.*.. *..*..* *** *****.*.**
CE YDDDPLDDKNAIITGNLLIVMAQIIVAIQMVYEQKYLTKYDVPALFAVGLEGLFGMVTLS
HP LLLVPMYYIPAG-SFSGNPRGTLEDALDAFCQVGQQPLIAVALLGNISSIAFFNFAGISV
   .*..*.***..  .**.** * ***.. *. .....* **.** *.. *********.**
CE ILMIPFYYIHVPRTFSTNPEGRLEDVFYAWKEITEEPTIALALSGTVVSIAFFNFAGVSV
HP TKELSATTRMVLDSLRTVVIWALSLALGWEAFHALQILGFLILLIGTALYNGLHRPLLGR
   **************.**.***..*..*  * * *.*. ** .*..** .**..
CE TKELSATTRMVLDSVRTLVIWVVSIPLFHEKFIAIQLSGFAMLILGTLIYNDILIGPWFR
HP LSRGRPLAEESEQERLLGGTRTPINDAS
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. N50907) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10488> (SEQ ID Nos. 67, 77, and 87)

Determination of the whole base sequence of the cDNA insert of clone HP10488 obtained from cDNA library of human liver revealed the structure consisting of a 39-bp 5'-untranslated region, a 273-bp ORF, and a 421-bp 3'-untranslated region. The ORF codes for a protein consisting of 90 amino acid residues and there existed one putative transmembrane domain at the N-terminus. Figure 27 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 10 kDa that was almost identical with the molecular weight of 10,151 predicted from the ORF. When expressed in COS7 cells, an expression product of about 6 kDa was observed in the membrane fraction.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. H73534) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein

<HP10538> (SEQ ID Nos. 68, 78, and 88)

Determination of the whole base sequence of the cDNA insert of clone HP10538 obtained from cDNA library of human osteosarcoma cell line Saos-2 revealed the structure consisting of a 357-bp 5'-untranslated region, a 1500-bp ORF, and a 1911-bp 3'-untranslated region. The ORF codes for a protein consisting of 499 amino acid residues and there existed at least four putative transmembrane domains. Figure 28 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of high molecular weight.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the mouse pore-forming $K^+$ channel subunit (GenBank Accession No. AF056492). Table 19 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the mouse pore-forming $K^+$ channel subunit (MM). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 32.4% in the N-terminal region of 241 amino acid residues.

Table 19

```
HP  MVDRGPLLTSAIIFYLAIGAAIFEVLEEPHWKEAKKNYYTQKLHLLKEFPCLGQEGLDK
            * . ...**. ** .*..**.** ..*.*.    . ..*.. **..*..*..
MM MRSTTLLALLALVLLYLVSGALVFQALEQPHEQQAQKKMDHGRDQFLRDHPCVSQKSLED
HP ILEVVSDAAGQG-----VAITGNQTFNNWNWPNAMIFAATVITTIGYGNVAPKTPAGRLF
   ..... .* * *        .....  ..**  .*..*..*.********.. .* *****
MM FIKLLVEALGGGANPETSWTNSSNHSSAWNLGSAFFFSGTIITTIGYGNIVLHTDAGRLF
HP CVFYGLFGVPLCLTWISALGKFFGGRAKR----LGQFLTKRGVSLRKAQITCTVIFIVWG
   *.**.* *.**    ....*. .*.. .*    ..... *. *.   ..  ..*.*.. *
MM CIFYALVGIPLFGMLLAGVGDRLGSSLRRGIGHIEAIFLKWHVPPGLVRSLSAVLFLLIG
HP VLVHLVIPPFVFMVTEGWNYIEGLYYSFITISTIGFGDFVAGVNPSANYHALYRYFVELW
    *. ...*.***   *.*. .*..*. ..*..*.****.*.* . ...  . *. .* .*
MM CLLFVLTPTFVFSYMESWSKLEAIYFVIVTLTTVGFGDYVPG-DGTGQNSPAYQPLVWFW
HP IYLGLAWLSLFVNWKVSMFVEVHKAIKKRRRRRKESFESSPHSRKALQVKGSTASKDVNI
   * .***...
MM ILFGLAYFASVLTTIGNWLRAVSRRTRAEMGGLTAQAASWTGTVTARVTQRTGPSAPPPE
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. R25184) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10542> (SEQ ID Nos. 69, 79, and 89)

Determination of the whole base sequence of the cDNA insert of clone HP10542 obtained from cDNA library of human [illegible] 5'-untranslated region, a 321-bp ORF, and a 426-bp 3'-

untranslated region. The ORF codes for a protein consisting of 106 amino acid residues and there existed one putative transmembrane domain. Figure 29 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 12 kDa that was almost identical with the molecular weight of 11,724 predicted from the ORF. When expressed in COS7 cells, an expression product of about 13 kDa was observed in the membrane fraction.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA029683) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10571> (SEQ ID Nos. 70, 80, and 90)

Determination of the whole base sequence of the cDNA insert of clone HP10571 obtained from cDNA library of human stomach cancer revealed the structure consisting of a 95-bp 5′-untranslated region, a 459-bp ORF, and a 675-bp 3′-untranslated region. The ORF codes for a protein consisting of 152 amino acid residues and there existed one putative transmembrane domain. Figure 30 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product

...,002 predicted from the ORF. In this case, the addition of a microsome led to the formation of a product of 23 kDa

which is considered to have a sugar chain being attached after secretion. In addition, there exists in the amino acid sequence of this protein one site at which N-glycosylation may occur (Asn-Ile-Thr at position 10).

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA105822) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP01470> (SEQ ID Nos. 91, 101, and 111)

Determination of the whole base sequence of the cDNA insert of clone HP01470 obtained from cDNA library of human stomach cancer revealed the structure consisting of a 157-bp 5′-untranslated region, a 1077-bp ORF, and a 385-bp 3′-untranslated region. The ORF codes for a protein consisting of 358 amino acid residues and there existed one putative transmembrane domain. Figure 31 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 43 kDa that was somewhat larger than the molecular weight of 40,489 predicted from the ORF. In this case, the addition of a microsome led to the formation of a product of 40 kDa from which the secretory signal is considered to have been cleaved and a product of 43.5 kDa which is considered to have been subjected to some modification. Application of the

the secretory signal sequence, allows to expect that the mature protein starts from glycine at position 23. When

expressed in COS7 cells, an expression product of about 44 kDa was observed in the supernatant fraction.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Caenorhabditis elegans hypothetical protein 39.9 kDa (SWISS-PROT Accession No. Q10005). Table 20 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the C. elegans hypothetical protein 39.9 kDa (CE). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 58.9% in the entire region.

Table 20

```
HP MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKLALQLHPDRNPDD
                             *.. *  ***********...*. ..********** .*******.**
CE  MRILNVSLLVLASSLVAFVECGRDFYKILGVAKNANANQIKKAYRKLAKELHPDRNQDD
HP PQAQEKFQDLGAAYEVLSDSEKRKQYDTYGEEGL--KDGHQSSHGDIFSHFFGDFGFMFG
     *.******..********.***   ** .****.  ..*   ..   * ** ***** *  *
CE EMANEKFQDLSSAYEVLSDKEKRAMYDRHGEEGVAKMGGGGGGGHDPFSSFFGDF-FG-G
HP GTPRQQDRNIPRGSDIIVDLEVTLEEVYAGNFVEVVRNKPVARQAPGKRKCNCRQEMRTT
   *.  .  . ..*.*.*...** *******.*.***. *.*.* .*..*.*.****.****.
CE GGGHGGEEGTPKGADVTIDLFVTLEEVYNGHFVEIKRKKAVYKQTSGTRQCNCRHEMRTE
HP QLGPGRFQMTQEVVCDECPNVKLVNEERTLEVEIEPGVRDGMEYPFIGEGEPHVDGEPGD
   *.*.***** *  ***********.*...****.* *. .* .   * ******..*.***
CE QMGQGRFQMFQVKVCDECPNVKLVQENKVLEVEVEVGADNGHQQIFHGEGEPHIEGDPGD
HP LRFRIKVVKHPIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVHISRDKITRPGAK
   *.*.*.. *** ***.************ ..* ****.* ***** *.. ***.*.***.
CE LKFKIRIQKHPRFERKGDDLYTNVTISLQDALNGFEMEIQHLDGHIVKVQRDKVTWPGAR
HP LWKKGEGLPNFDNNNIKGSLIITFDVDFPKEQLTEEAREGIKQLLKQGSVQ-KVYNGLQG
   *.**.**.*....** ** *..****.***..*..*... * ..*.*..*. *.****
CE LRKKDEGMPSLEDNNKKGMLVVTFDVEFPKTELSDEQKAQIIEILQQNTVKPKAYNGL
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA282838) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP02419> (SEQ ID Nos. 92, 102, and 112)

[illegible] insert of clone HP02419 obtained from cDNA library of human stomach cancer revealed the structure consisting of a 253-bp

5'-untranslated region, a 681-bp ORF, and a 1120-bp 3'-untranslated region. The ORF codes for a protein consisting of 226 amino acid residues and there existed four putative transmembrane domains. Figure 32 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of high molecular weight.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the human hypothetical protein KIAA0108 (SWISS-PROT Accession No. Q15012). Table 21 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the human hypothetical protein KIAA0108 (KI). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 43.9% in the entire region.

Table 21

```
HP     MKMVAPWTRFYSNSCCLCCHVRTGTILLGVWYLIINAVVLLILLSALADPD---QY
               ****..** *********.**.**...* .. ..*  ....*.    .
KI MVSMSFKRNRSDRFYSTRCCGCCHVRTGTIILGTWYMVVNLLMAILLTVEVTHPNSMPAV
HP NFSSSELGGDFEF-MDDANMCIAIAISLLMILICAMATYGAYKQRAAWIIPFFCYQIFDF
   *.  . .*. ..  . ..* *. .*.*.**..*..* .*** . ...*.******..***
KI NIQYEVIGNYYSSERMADNACVLFAVSVLMFIISSMLVYGAISYQVGWLIPFFCYRLFDF
HP ALNMLVAITVLIYPNSIQEYIRQLPPNFPYRDDVMSVNPTCLVLIILLFISIILTFKGYL
   .*. ****. *.*  .*.**. ** *.***.**.......**..*.*.*......**.**
KI VLSCLVAISSLTYLPRIKEYLDQL-PDFPYKDDLLALDSSCLLFIVLVFFALFIIFKAYL
HP ISCVWNCYRYINGRNSSDVLVYVT-SNDTTVLLPPYDDATVNGAAKEPPPPYVSA
   *.******.***.** ... ** . .. .  .**.* . .*. ..*******..*
KI INCVWNCYKYINNRNVPEIAVYPAFEAPPQYVLPTY-EMAVKMPEKEPPPPYLPA
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA173214) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP02631> (SEQ ID Nos. 93, 103, and 113)

Determination of the whole base sequence of the cDNA insert of clone HP02631 obtained from cDNA library of human osteosarcoma cell line Saos-2 revealed the structure consisting of a 42-bp 5'-untranslated region, a 588-bp ORF,

[illegible]

this codon encodes selenocysteine from the molecu[illegible]ght

of the translation product and the sequence comparison data with the Caenorhabditis elegans homologue. The ORF codes for a protein consisting of 195 amino acid residues and there existed a putative secretory signal at the N-terminus and one putative transmembrane domain in the intermediate region. Figure 33 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 58 kDa. In this case, the addition of a microsome led to the formation of a product of 56 kDa from which the secretory signal is considered to have been cleaved. Since both of these products are larger than the molecular weight of 22 kDa predicted from the ORF, it is likely that the protein interacts with another protein.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Caenorhabditis elegans hypothetical protein C35C5.3 (EMBL Accession No. Z78417). Table 22 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the C. elegans hypothetical protein C35C5.3 (CE). U at position 49 in the amino acid sequence of the protein of the present invention represents selenocysteine. Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 37.9% in the entire region other than the N-terminal region. Cystein was found in the sequence of the C.

elegans ...

49 encoded by the stop codon (selenocysteine) ... of the present invention.

Table 22

```
HP                                                   MRLLLL

CE MRIHDELQKQDMSRFGVFIIGVLFFMSVCDVLRTEEHSHDENHVHEKDDFEAEFGDETDS
HP LLVAASAMVRSEASANLGGVPSKRLKMQYATGPLLKFQICVSUGYRRVFEEYMRVISQRY
                                    * *..  *** **...*...   ...*

CE QSFSQGTEEDHIEVREQSSFVKPTAVHHAKDLPTLRIFYCVSCGYKQAFDQFTTFAKEKY
HP PDIRIEGENYLPQPIYRHIASFLSVFKLVLIGLIIVGKDPFAFFGMQAPSIWQWGQENKV
   *...***.*. *     ..*  **  *.... *.. * .**. **.  * * * ...**.

CE PNMPIEGANFAPVLWKAYVAQALSFVKMAVLVLVLGGINPFERFGLGYPQILQHAHGNKM
HP YACMMVFFLSNMIENQCMSTGAFEITLNDVPVWSKLESGHLPSMQQLVQILDNEMKLNVH
    .**.**.*.*..*.  .******. *.. ..***.***..** *...*..*... .

CE SSCMLVFMLGNLVEQSLISTGAFEVYLGNEQIWSKIESGRVPSPQEFMQLIDAQLAVLGK
HP MDSIPHHRS

CE APVNTESFGEFQQTV
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA156969) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP02695> (SEQ ID Nos. 94, 104, and 114)

Determination of the whole base sequence of the cDNA [illegible] cDNA library of human [illegible] 5'-untranslated region, a 1020-bp ORF, and a 160-bp 3'-

untranslated region. The ORF codes for a protein consisting of 339 amino acid residues and there existed three putative transmembrane domains. Figure 34 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 38 kDa that was almost identical with the molecular weight of 38,274 kDa predicted from the ORF.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the rat hypertension-induced protein S-2 fragment (PIR Accession No. 539959). Table 23 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the rat hypertension-induced protein S-2 fragment (RN). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 74.3% in the entire region.

Table 23

```
HP MNWELLLWLLVLCALLLLLVQLLRFLRADGDLTLLWAEWQGRRPEWELTDMVVWVTGASS

HP GIGEELAYQLSKLGVSLVLSARRVHELERVKRRCLENGNLKEKDILVLPLDLTDTGSHEA
                             ****.******************.**.**.
RN                           VKRRSLENGNLKEKDILVLPLDLADTSSHDI
HP ATKAVLQEFGRIDILVNNGGMSQRSLCMDTSLDVYRKLIELNYLGTVSLTKCVL[illegible]MIER
   ***.*****************... **  .*..*... ***.*********** ****..*
RN ATKTVLQEFGRIDILVNNGGVAHASLVENTNMDIFKVLIEVNYLGTVSLTKCFLPHMMER
HP KQGKIVTVNSILGIISVPLSIGYCASKHALRGFFNGLRTELATYPGIIVSNICPGPVQSN
   .*****...*
RN NQGKIVVMKS
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. T84331) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10031> (SEQ ID Nos. 95, 105, and 115)

Determination of the whole base sequence of the cDNA insert of clone HP10031 obtained from cDNA library of human osteosarcoma cell line Saos-2 revealed the structure consisting of a 55-bp 5'-untranslated region, a 1464-bp ORF, and a 649-bp 3'-untranslated region. The ORF codes for a [illegible]

depicts the hydrophobicity/hydrophilicity pr[illegible]

by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of high molecular weight. When expressed in COS7 cells, an expression product of about 55 kDa was observed in the membrane fraction.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Caenorhabditis elegans hypothetical protein CELK07H8 (GenBank Accession No. AF047659). Table 24 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the C. elegans hypothetical protein CELK07H8 (CE). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 44.2% in the entire region.

Table 24

```
HP                                  MDGTETRQRRLDSCGKPGELGLPHPLSTGGLPVAS

CE MKGGGGIGDGKKDYQSAVHEGLTTFDQLGIALEDVGKSMDAETATPGGSLFSRVIFRFRN
HP EDGALRAPESQSVTPKPLETEPSRETAWSIGLQVTVPFMFAGLGLSWAGMLLDYFQHWPV
   *...*.... .  .        .  . *... . .**   **   ****. .**..*.    **.
CE ENSSLKSRTYDHSNDLVNMSVIPAESSYVLFFQVLFPFAVAGLGMVFAGLVLSIVVTWPL
HP FVEVKDLLTLVPPLVGLKGNLEMTLASRLSTAANTGQIDDPQEQHRVISSNLALIQVQAT
   * *. ..*.***.*.**************** ** *..*...... *. .****.*****
CE FEEIPEILILVPALLGLKGNLEMTLASRLSTLANLGHMDSSKQRKDVVIANLALVQVQAT
HP VVGLLAAVAALLLGVVSREEVDVAKVELLCASSVLTAFLAAFALGVLMVCIVIGARKLGV
   **..**.. *  *. ..... * *. .*.****. **  *...*..*** .....** ..
CE VVAFLASAFAAALAFIPSGDFDWAHGALMCASSLATACSASLVLSLLMVVVIVTSRKYNI
HP NPDNIATPIAASLGDLITLSILALVSSFFYR-HKDSRYLTPLVCLSFAALTPVWVLIAKQ
   ****.***********.**..**. .* * . *.....*. .* . *  * * *. **..
CE NPDNVATPIAASLGDLTTLTVLAFFGSVFLKAHNTESWLNVIVIVLFLLLLPFWIKIANE
HP SPPIVKILKFGWFPIILAMVISSFGGLILSKTVSKQQYKGMAIFTPVICGVGGNLVAIQT
   .  . ..*  ** *.*..*.*** **.**...*  ..*.......**. ******.*.*.
CE NEGTQETLYNGWTPVIMSMLISSAGGFILETAV--RRYHSLSTYGPVLNGVGGNLAAVQA
HP SRISTYLHMWSAPGVLPLQ--MKKFWPNPCSTFCTSEINSMSARVLLLLVVPGHLIF-FY
   **.***.*. .. ****  .   ...* ..  ..* ..* .*.*************. * *
CE SRLSTYFHKAGTVGVLPNEWTVSRF-TSVQRAFFSKEWDSRSARVLLLLVVPGHICFNFL
HP I-IYLVEGQSVINSQ--TFVVLYLLAGLIQVTILLYLAEVMVRLTWHQALDPDNHCIPYL
   * .. ..... ....    *. **..*..***.***.. ...* * *.   .****  ****
CE IQLFTLTSKNNVTPHGPLFTSLYMIAAIIQVVILLFVCQLLVALLWKWKIDPDNSVIPYL
HP TGLGDLLGTGLLALCFFTDWLLKSKAELGGISELASGPP
   *.********** . *.*
CE TALGDLLGTGLLFIVFLTTDHFDPKELTSS
```

[illegible]

sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for

example, Accession No. AA334000) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10530> (SEQ ID Nos. 96, 106, and 116)

Determination of the whole base sequence of the cDNA insert of clone HP10530 obtained from cDNA library of human osteosarcoma cell line Saos-2 revealed the structure consisting of a 80-bp 5'-untranslated region, a 1182-bp ORF, and a 95-bp 3'-untranslated region. The ORF codes for a protein consisting of 393 amino acid residues and there existed a putative secretory signal at the N-terminus. Figure 36 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 46 kDa that was somewhat larger than the molecular weight of 44,912 predicted from the ORF. In this case, the addition of a microsome led to the formation of a product of 45.5 kDa from which the secretory signal is considered to have been cleaved. Application of the (-3,-1) rule, a method for predicting the cleavage site of the secretory signal sequence, allows to expect that the mature protein starts from lysine at position 23. When expressed in COS7 cells, an expression product of about 43 kDa was observed in the supernatant fraction and the membrane fraction.

The search of the protein data base using the amino acid sequence of the present protein revealed that the [illegible] protein 1G002N01 (GenBank Accession No. AF007269). Table 25 shows the comparison between amino acid sequences of the

human protein of the present invention (HP) and the A. thaliana hypothetical protein IG002N01 (AT). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 27.0% in the N-terminal region of 355 amino acid residues.

Table 25

```
HP                                                             MRTLFNLLWL

AT MELTSFQKSPSSNDVVSFSVSLVRNSMARRRRSSAAESLKRRNDGYESLCQVVQQDSDRR
HP ALACSPVHTTLSKSDAKKAASKTLLEKSQFSDKPVQDRGLVVTDLKAESVVLEHRSYCSA
                                    ........*.* **.. .... **..  ..
AT LITIFVIFFIVIPAVSIAVYKVKFADRVIQTESSIRQKGIVKTDINFQEILTEHSK--AS
HP KARDRHFAGDVLGYVTPWNSHGYDVTKVFGSKFTQISPVWLQ-LKRRGREMFEVTGLHDV
   ....**..  **.*.**  ..* ..  *....  . .  *.* *..**... . .**...
AT ENSTRHYDYPVLAYITP--CQGSGL--VLEGR-HNADKGWIQELRSRGNALSASKGLPKL
HP DQGWMRAVRKHAKGLHIVPRLLFEDWTYDDFRNVLDSEDEIEELSKTVVQVAKNQHFDGF
      . . . .  * ... .  * .  **  .. .*.* *.*  . ..   .. . .. .
AT ---YNSCIFHALKRMNFFTLELVNFNTYLVIMFALNS-REMEYNGIVLESWSRWAAYGVL
HP VVEVWNQLLSQKRVGLIHMLTHLAEALHQARLLALLVIPPAITPGTDQLGMFTHKEFEQL
        ...  .    * . *  ... .... .... .  .  ... *.  *...... *
AT HDPDLRKMALKFVKQLGDALHSTSSPRNNQQHMQFMYVVGPPRSEKLQMYDFGPEDLQFL
HP APVLDGFSLMTYDYSTAHQPGPNAPLSWVRACVQ-VLDPKSK----WRSKILLGLNFYGM
      .*********.*....******..*.   .. .*.....      .*.***.****
AT KDSVDGFSLMTYDFSNPQNPGPNAPVKWIDLTLKLLLGSSNNIDSNIARKVLLGINFYGN
HP DYATSKDAREPVVGARYIQTLKDHRPRMVWDSQASEHFFEYKKSRSGRHVVFYPTLKSLQ
   *...* .. ....*  *.. *..*.* . **....**.* *..... .*.******.*.
AT DFVISGGGGGAITGRDYLALLQKHKPTFRWDKESGEHLFMYRDDKNIKHAVFYPTLMSIL
HP VRLELARELGVGVSIWELGQGLDYFYDLL
   .*** **  *.*.****.**. ..*
AT LRLENARLWGIGISIWEIGQDKGHFGKYAEASLEASSIFSGHTFDMQFRTNPRQLSRNGS
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA302913) in ESTs, but since they [illegible] any of these sequences codes for the same protein as the

protein of the present invention.

<HP10541> (SEQ ID Nos. 97, 107, and 117)

Determination of the whole base sequence of the cDNA insert of clone HP10541 obtained from cDNA library of human stomach cancer revealed the structure consisting of a 7-bp 5'-untranslated region, a 591-bp ORF, and a 113-bp 3'-untranslated region. The ORF codes for a protein consisting of 196 amino acid residues and there existed a putative secretory signal at the N-terminus. Figure 37 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 23 kDa that was somewhat larger than the molecular weight of 21,553 predicted from the ORF. In this case, the addition of a microsome led to the formation of a product of 20 kDa from which the secretory signal is considered to have been cleaved and a product of 23 kDa which is considered to have a sugar chain being attached. Application of the (-3,-1) rule, a method for predicting the cleavage site of the secretory signal sequence, allows to expect that the mature protein starts from glycine at position 41. In addition, there exists in the amino acid sequence of this protein one site at which N-glycosylation may occur (Asn-Leu-Thr at position 185).

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the human zymogen membrane protein (GenBank Accession No. AF056492). Table 26 shows the [illegible] of the present invention (HP) and the human zymogen membrane protein (ZM). Therein, the marks of -, *, and . represent a

gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 37.6% in the C-terminal region of 133 amino acid residues.

Table 26

```
HP MWRVPGTTRRPVTGESPGMHRPEAMLLLLTLALLGGPTWAGKMYGPGGGKYFS-TTEDYD
                                              **.**** ** .... *
ZM               MLTVALLALLCASASGNAIQARSSSYSGEYGSGGGKRFSHSGNQLD
HP HEITGLRVSVGLLLVKSVQVKLGDSWDVKLGALGGNTQEVTLQPGEYITKVFVAFQAFLR
     **.***.*.   . ..**. *. *.  .*. .*. .*. *.*** ...*   .. .*.
ZM GPITALRVRVNTYYIVGLQVRYGKVWSDYVGGRNGDLEEIFLHPGESVIQVSGKYKWYLK
HP GMVMYTSKDRYFYFGKLDGQISSAYPSQEGQVLVGIYGQYQLLGIKSIGFEWN-YPLEEP
    .*. *.*.**. *** .*   .* * . . **  * *.   * *..**..*. **
ZM KLVFVTDKGRYLSFGKDSGTSFNAVPLHPNTVLRFISGRSGSL-IDAIGLHWDVYPTSCS
HP TTEPPVNLTYSANSPVGR

ZM RC
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA340605) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

Determination of the whole base sequence of the cDNA

insert of clone HP10550 obtained from cDNA library of human stomach cancer revealed the structure consisting of a 241-bp 5′-untranslated region, a 324-bp ORF, and a 86-bp 3′-untranslated region. The ORF codes for a protein consisting of 107 amino acid residues and there existed one putative transmembrane domain. Figure 38 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of high molecular weight.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA348310) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10590> (SEQ ID Nos. 99, 109, and 119)

Determination of the whole base sequence of the cDNA insert of clone HP10590 obtained from cDNA library of human fibrosarcoma cell line HT-1080 revealed the structure consisting of a 77-bp 5′-untranslated region, a 1053-bp ORF, and a 180-bp 3′-untranslated region. The ORF codes for a protein consisting of 350 amino acid residues and there existed one putative transmembrane domain. Figure 39 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product

weight of [illegible] predicted from the ORF. In this case, the addition of a microsome led to the formation of a product of

43 kDa which is considered to have a sugar chain being attached. In addition, there exist in the amino acid sequence of this protein two sites at which N-glycosylation may occur (Asn-Asn-Ser at position 144 and Asn-Leu-Thr at position 328).

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA461346) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10591> (SEQ ID Nos. 100, 110, and 120)

Determination of the whole base sequence of the cDNA insert of clone HP10591 obtained from cDNA library of human fibrosarcoma cell line HT-1080 revealed the structure consisting of a 232-bp 5'-untranslated region, a 324-bp ORF, and a 844-bp 3'-untranslated region. The ORF codes for a protein consisting of 107 amino acid residues and there existed one putative transmembrane domain. Figure 40 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 12 kDa that was almost identical with the molecular weight of 11,328 predicted from the ORF.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for [illegible] partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein

of the present invention.

<HP01462> (SEQ ID Nos. 121, 131, and 141)

Determination of the whole base sequence of the cDNA insert of clone HP01462 obtained from cDNA library of human fibrosarcoma cell line HT-1080 revealed the structure consisting of a 121-bp 5'-untranslated region, a 1452-bp ORF, and a 477-bp 3'-untranslated region. The ORF codes for a protein consisting of 483 amino acid residues and there existed a putative secretory signal at the N-terminus. Figure 41 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 72 kDa that was larger than the molecular weight of 55,838 predicted from the ORF. Application of the (-3,-1) rule, a method for predicting the cleavage site of the secretory signal sequence, allows to expect that the mature protein starts from lysine at position 21.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Caenorhabditis elegans hypothetical protein ZK1058.4 (EMBL Accession No. Z35604). Table 27 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the C. elegans hypothetical protein ZK1058.4 (CE). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of

Table 27

```
HP MKAFHTFCVVLLVFGSVSEAKFDDFEDEEDIVEYDDNDFAEFEDVMEDSVTESPQRVIIT
                                                            * *
CE                                           MKIVWIFLIFFIGFAIST
HP EDDE-DETTVELEGQDENQEGDFEDADTQEGDTESEPYDDEEFEGYEDKP---------D
   .*.* .* . *. * ...*. ... ....   .... ...*.*.  *..*         *
CE DDNEFAEFEDEFVGSSATQAPEIQREGEPPVLKQKDDFEEEDFGVVEEEPEEAEKVREAD
HP TSSSKNKDPITIVDVPAHLQNSWESYYLEILMVTGLLAYIMNYIIGKNKNSRLAQAWFNT
   ....   .*....*****....*.** .* ..* .* *.  **.***..*...**. *.
CE SDDAAPAQPLKFADVPAHFRSNWASYQVEGIVVLIILIYMTNYLIGKTTNASIAQTIFDM
HP HRELLESNFTLVGDDGTNKEATSTGKLNQENEHIYNLWCSGRVCCEGMLIQLRFLKRQDL
    *  **..*..******.. .. . .*..... ... **.***  .....*....****.
CE CRPTLEEQFAVVGDDGTTDLDKMIPSLKHDTDSTFSAWCTGRVNVNSLFLQMKMVKRQDV
HP LNVLARMMRPVSDQVQIKVTMN-DEDMDTYVFAVGTRKALVRLQKEMQDLSEFCSDKPKS
   .. .  *. * .*.. **.... ..* *. .****..*   .  *** **..* *.. ..
CE VSRIMEMFTPSGDKMTIKASLETTNDTDPLIFAVGEKKIASKYFKEMLDLNSFASERKQA
HP GAKYGLPDSLAILSEMGEVTDGMMDTKMVHFLTHYADKIESVHFSDQFSGPKIMQEEGQP
   .....**.* .. .. .**. ...*. .* .*....* ** .*.****.***  ..*. .
CE AQQFNLPASWQVYADQNEVVFSILDPGVVSLLKKHEDAIEFIHISDQFTGPKPAEGESYT
HP LKLPDTKRTLLFTFNVPGSGNTYPKDMEALLPLMNMVIYSIDKAKKFRLNREGKQKADKN
    .**...* .. ..*..  *    .* *... ..*.*.* ****.*..*....* **...
CE -RLPEAQRYMFVSLNLQYLG----QDEESVMEILNLVFYLIDKARKMKLSKDAKVKAERR
HP RARVEENFLKLTHVQRQEAAQSRREEKKRAEKERIMNEEDPEKQRRLEEAALRREQKKLE
   *   *..*** **  ******.*****.*. *...*.*.***.*.***. .*.**.*
CE RKEFEDAFLKQTHQFRQEAAQARREEKTRERKQKLMDESDPERQKRLEAKELKREAKA--
HP KKQMKMKQIKVKAM
    *  ****.***
CE -KSPKMKQLKVK
```

sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for

example, Accession No. AA307793) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP02485> (SEQ ID Nos. 122, 132, and 142)

Determination of the whole base sequence of the cDNA insert of clone HP02485 obtained from cDNA library of human stomach cancer revealed the structure consisting of a 69-bp 5'-untranslated region, a 1005-bp ORF, and a 1672-bp 3'-untranslated region. The ORF codes for a protein consisting of 334 amino acid residues and there existed one putative transmembrane domain. Figure 42 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 36 kDa that was almost identical with the molecular weight of 38,171 predicted from the ORF. When expressed in COS7 cells, an expression product of about 23 kDa was observed in the membrane fraction.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Caenorhabditis elegans hypothetical protein W01A11.2 (GenBank Accession No. U64852). Table 28 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the C. elegans hypothetical protein W01A11.2 (CE). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention,

the present invention, respectively. The both proteins shared a homology of 45.5% in the entire region.

Table 28

```
HP                                 MVEFAPLFMPWERRLQTLAVLQFVFSFLALAEICT-V
                                     .***..**.***.****. *.* .. *. .   *

CE MRLRLSSISGKAKLPDKEICSSVSRILAPLLVPWKRRLETLAVMGFIFMWVILPIMDLWV
HP GFIALLFTRFWLLTVLYAAWWYLDRDKPRQGGRHIQAIRCWTIWKYMKDYFPISLVKTAE
    * .*. **.*.*. ***.*.* * *.*....*. .   *   . ***. .***..*.***.

CE PFHVLFNTRWWFLVPLYAVWFYYDFDTPKKASRRWNWARRHVAWKYFASYFPLRLIKTAD
HP LDPSRNYIAGFHPHGVLAVGAFANLCTESTGFSSIFPGIRPHLMMLTLWFRAPFFRDYIM
    * ..**** * ****...**.*....*..***.. ****..*.* *.  *  ** *.. .

CE LPADRNYIIGSHPHGMFSVGGFTAMSTNATGFEDKFPGIKSHIMTLNGQFYFPFRREFGI
HP SAGLVTSEKESAAHILNRKGGGNLLGIIVGGAQEALDARPGSFTLLLRNRKGFVRLALTH
     * .. .*** ...*.. * *.  .*..*** ***.*.*.. ** * **.** . **.

CE MLGGIEVSKESLEYTLTKCGKGRACAIVIGGASEALEAHPNKNTLTLINRRGFCKYALKF
HP GAPLVPIFSFGENDLFDQIPNSSGSWLRYIQNRLQKIMGISLPLFHGRGVF-QYSFGLIP
   ** ***...******..*  *..**.** .*.......*.. **..**..* ** .**.*

CE GADLVPMYNFGENDLYEQYENPKGSRLREVQEKIKDMFGLCPPLLRGRSLFNQYLIGLLP
HP YRRPITTVVGKPIEVQKTLHPSEEEVNQLHQRYIKELCNLFEAHKLKFNIPADQHLEFC
   .*.*.***.*.** * .* .*. *....**..* ..*.****..*   .**.* ** *

CE FRKPVTTVMGRPIRVTQTDEPTVEQIDELHAKYCDALYNLFEEYKHLHSIPPDTHLIFQ
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. D25664) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

[illegible] (SEQ ID NOS. [illegible]

Determination of the whole base sequence of the cDNA

insert of clone HP02798 obtained from cDNA library of human fibrosarcoma cell line HT-1080 revealed the structure consisting of a 31-bp 5'-untranslated region, a 804-bp ORF, and a 301-bp 3'-untranslated region. The ORF codes for a protein consisting of 267 amino acid residues and there existed four putative transmembrane domains. Figure 43 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 29 kDa that was almost identical with the molecular weight of 30,778 predicted from the ORF. When expressed in COS7 cells, an expression product of about 26 kDa was observed in the membrane fraction.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the human DHHC-containing cysteine-rich protein (GenBank Accession No. U90653). Table 29 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the human DHHC-containing cysteine-rich protein (DH). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 35.0% in the intermediate region of 100 amino acid residues. The positions of seven cysteines were conserved between the two proteins. The protein of the present invention also had the DHHC (Asp-His-His-Cys) sequence.

Table 29

```
HP                               MAPWALLSPGVLVRTGHTVLTWGI

DH MYKMNICNKPSNKTAPEKSVWTAPAQPSGPSPELQGQRSRRNGWSWPPHPLQIVAWLLYL
HP TLVLFLHDTELRQWEEQGELLLPLTFLLLVLGSLLLYLAVSLMDPGYVNVQPQP-QEELK
                              * *...*..  .**.  **. .. .. *

DH FFAVIGFGILVPLLPHHWVPAGYACMGAIFAGHLVVHLTAVSIDPADDNVRDKSYAGPLP
HP EEQTAMVPPAIPLRRCRYCLVLQPLRARHCRECRRCVRRYDHHCPWMENCVGERNHPLFV
     . .  ...*   .*. * *  . *..**..*..**  .**** *..*******..**.

DH IFNRSQHAHVIEDLHCNLCNVDVSARSKHCSACNKCVCGFDHHCKWLNNCVGERNYRLFL
HP VYLALQLVVLLWGLYLAWSGLRFFQPWGLWLRSSGLLFATFLLLSLFSLVASLLLVSHLY
     .* .*. .*

DH HSVASALLGVLLLVLGGHICLRGVLCQPHASAHQPTL
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. D79050) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10041> (SEQ ID Nos. 124, 134, and 144)

Determination of the whole base sequence of the cDNA insert of clone HP10041 obtained from cDNA library of human osteosarcoma cell line Saos-2 revealed the structure consisting of a 12-bp 5′-untranslated region, a 321-bp ORF, [illegible] The ORF codes for [illegible]

[illegible]

existed one putative transmembrane domain. Figure 44 depicts

the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 12 kDa that was almost identical with the molecular weight of 12,060 predicted from the ORF. When expressed in COS7 cells, an expression product of about 13 kDa was observed in the membrane fraction.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Caenorhabditis elegans hypothetical protein K10B2.4 (GenBank Accession No. U28730). Table 30 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the C. elegans hypothetical protein K10B2.4 (CE). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 62.1% in the entire region.

Table 30

```
HP MSTNNMSDPRRPNKVLRYKP---PPSECNPALDDPTPDYMNLLGMIFSMCGLM[illegible]KWCA
           .****.*...****    .... ..   .** *.***.************    **.
CE   MQQNGDPRRTNRIVRYKPLDSTANQQQAISEDPLPEYMNVLGMIFSMCGLM[illegible]
HP WVAVYCSFISFANSRSSEDTKQMMSSFMLSISAVVMSYLQNPQPMTPPW
   *.*.  **  *****.*.*.*.**..******.*********** *..***
CE WLALVCSCISFANTRTSDDAKQIVSSFMLSVSAVVMSYLQNPSPIIPPWVT[illegible]
```

Furthermore, the search of the GenBank [illegible]

sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. H20098) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10246> (SEQ ID Nos. 125, 135, and 145)

Determination of the whole base sequence of the cDNA insert of clone HP10246 obtained from cDNA library of human epidermoid carcinoma cell line KB revealed the structure consisting of a 110-bp 5′-untranslated region, a 675-bp ORF, and a 79-bp 3′-untranslated region. The ORF codes for a protein consisting of 224 amino acid residues and there existed five putative transmembrane domains. Figure 45 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 23 kDa that was somewhat smaller than the molecular weight of 25,244 predicted from the ORF. When expressed in COS7 cells, an expression product of about 21 kDa was observed in the membrane fraction.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the human putative seven transmembrane domain protein (GenBank Accession No. Y18007). Table 31 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the human putative seven transmembrane domain protein (TM). [illegible] [illegible] residue identical with that of the protein of the present invention, and an amino acid residue similar to that

of the protein of the present invention, respectively. The both proteins shared a homology of 93.3% in the entire region.

Table 31

```
HP MTLFHFGNCFALAYFPYFITYKCSGLSEYNAFWKCVQAGVTYLFVQLCKMLFLATFFPTW
   ***********************..***********************************
TM MTLFHFGNCFALAYFPYFITYKCTDLSEYNAFWKCVQAGVTYLFVQLCKMLFLATFFPTW
HP EGGIYDFIGEFMKASVDVADLIGLNLVMSRNAGKGEYKIMVAALGWATAELIMSRCIPLW
   ************************************************************
TM EGGIYDFIGEFMKASVDVADLIGLNLVMSRNAGKGEYKIMVAALGWATAELIMSRCIPLW
HP VGARGIEFDWKYIQMSIDSNISLVHYIVASAQVWMITRYDLYHTFRPAVLLLMFLSVYKA
   *********************** .******************************.****
TM VGARGIEFDWKYIQMSIDSNISLGPYIVASAQVWMITRYDLYHTFRPAVLLLMFLRVYKA
HP FVMETFVHLCSLGSWAALLARAVVTGLLALSTLALYVAVVNVHS
   ****************.*.* .**.***....**.*********
TM FVMETFVHLCSLGSWAVLMAGVVVKGLLVIRNLAMYVAVVNVHS
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA453931) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10392> (SEQ ID Nos. 126, 136, and 146)

insert of clone HP10392 obtained from cDNA library of human osteosarcoma cell line U-2 OS revealed the structure

consisting of a 24-bp 5′-untranslated region, a 777-bp ORF, and a 726-bp 3′-untranslated region. The ORF codes for a protein consisting of 258 amino acid residues and there existed a putative secretory signal at the N-terminus. Figure 46 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 34 kDa that was somewhat larger than the molecular weight of 29,623 predicted from the ORF. Application of the (-3,-1) rule, a method for predicting the cleavage site of the secretory signal sequence, allows to expect that the mature protein starts from leucine at position 49.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. H15999) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention. In addition, partial identity with the hypothetical protein KIAA0384 (Accession No. AB002382) was observed, although the hypothetical protein had a different ORF.

<HP10489> (SEQ ID Nos. 127, 137, and 147)

Determination of the whole base sequence of the cDNA insert of clone HP10489 obtained from cDNA library of human [illegible] the structure consisting of a 137-bp [illegible] untranslated region. The ORF codes for a [illegible] of 110 amino acid residues and there existed two putative transmembrane domains. Figure 47 depicts the

hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 19 kDa that was somewhat larger than the molecular weight of 12,010 predicted from the ORF.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA262162) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10519> (SEQ ID Nos. 128, 138, and 148)

Determination of the whole base sequence of the cDNA insert of clone HP10519 obtained from cDNA library of human stomach cancer revealed the structure consisting of a 67-bp 5'-untranslated region, a 276-bp ORF, and a 367-bp 3'-untranslated region. The ORF codes for a protein consisting of 91 amino acid residues and there existed one putative transmembrane domain. Figure 48 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 10 kDa that was almost identical with the molecular weight of 10,275 predicted from the ORF.

[illegible]thermore, the search of the GenBank using the base [illegible] has revealed the registration [illegible] more (for [illegible] of sequences that [illegible] example, Accession No. W16639) in ESTs, [illegible] partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein

of the present invention.

<HP10531> (SEQ ID Nos. 129, 139, and 149)

Determination of the whole base sequence of the cDNA insert of clone HP10531 obtained from cDNA library of human osteosarcoma cell line Saos-2 revealed the structure consisting of a 55-bp 5'-untranslated region, a 1035-bp ORF, and a 1092-bp 3'-untranslated region. The ORF codes for a protein consisting of 344 amino acid residues and there existed five putative transmembrane domains. Figure 49 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of high molecular weight.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. R50695) in ESTs, but, since they are partial sequences, it can not be judged whether or not any of these sequences codes for the same protein as the protein of the present invention.

<HP10574> (SEQ ID Nos. 130, 140, and 150)

Determination of the whole base sequence of the cDNA insert of clone HP10574 obtained from cDNA library of human stomach cancer revealed the structure consisting of a 210-bp 5'-untranslated region, a 1287-bp ORF, and a 1276-bp 3'-untranslated region. The ORF codes for a protein consisting of 428 amino acid [illegible] at the N-terminus and one [illegible] secretory signal at the N-terminus and one transmembrane domain in the intermediate region. Figure 50 depicts the hydrophobicity/hydrophilicity profile, obtained

by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of high molecular weight. Application of the (-3,-1) rule, a method for predicting the cleavage site of the secretory signal sequence, allows to expect that the mature protein starts from serine at position 36.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Drosophila melanogaster GOLIATH protein (SWISS-PROT Accession No. Q06003). Table 32 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the D. melanogaster GOLIATH protein (DM). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The intermediate region of 169 amino acids of the protein of the present invention shared a homology of 41.4% with the N-terminal region of the D. melanogaster GOLIATH protein.

Table 32

```
HP MGPPPGAGVSCRGGCGFSRLLAWCFLLALSPQAPGSRGAEAVWTAYLNVSWRVPHTGVNR
HP TVWELSEEGVYGQDSPLEPVAGVLVPPDGPGALNACNPHTNFTVPTVWGSTVQVSWLALI
HP QRGGGCTFADKIHLAYERGASGAVIFNFPGTRNEVIPMSHPGAVDIVAIMIGNLKGTKIL
                                              .*.*...  .  * ..
                                MQLEKMQIKGKTRNIAAVITYQNIGQDLS
DM
HP QSIQRGIQVTMVIEVGKK---HGPWVNHYSIFFVSVSFFIITAATVGYFIFYSARRLRNA
     ....* .**. *  *..      . .*. *..***.* **. .... ..***  .*.*
DM LTLDKGYNVTISIIEGRRGVRTISSLNRTSVLFVSIS-FIV-DDILCWLIFYYIQRFRYM
HP RAQSRKQRQLKADAKKAIGRLQLRTLKQGDKEIGPDGDSCAVCIELYKPNDLVRILTCNH
    .*.... *.* . .****  ... .* * .* * . *.*.**.*** ***.* .***.*.*
DM QAKDQQSRNLCSVTKKAIMKIPTKTGKFSD-EKDLDSDCCAICIEAYKPTDTIRILPCKH
HP IFHKTCVDPWLLEHRTCPMCKCDILKALGIEVDVEDGSVSLQVPVSNEISNSASSHEEDN
    ***.*.****.********* *.**  *  *.
DM EFHKNCIDPWLIEHRTCPMCKLDVLKFYGYVVGDQIYQTPSPQHTAPIASIEEVPVIVVA
HP RSETASSGYASVQGTDEPPLEEHVQSTNESLQLVNHEANSVAVDVIPHVDNPTFEEDETP

DM VPHGPQPLQPLQASNMSSFAPSHYFQSSRSPSSSVQQQLAPLTYQPHPQQAASERGRRNS
HP NQETAVREIKS

DM APATMPHAITASHQVTDV
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA155685) in ESTs, but, since they [illegible] sequences, it can not be judged whether or not [illegible] protein of the present invention.

INDUSTRIAL APPLICABILITY

The present invention provides human proteins having hydrophobic domains, DNAs coding for these proteins, and expression vectors for these DNAs as well as eucaryotic cells expressing these DNAs. All of the proteins of the present invention are secreted or exist in the cell membrane, so that they are considered to be proteins controlling the proliferation and/or the differentiation of the cells. Accordingly, the proteins of [illegible] present invention can be employed as pharmaceuticals such as carcinostatic agents which act to control the proliferation and/or the differentiation of the cells, or as antigens for preparing antibodies against these proteins. The DNAs of the present invention can be utilized as probes for the genetic diagnosis and gene sources for the gene therapy. Furthermore, the DNAs can be utilized for large-scale expression of these proteins. Cells into which these genes are introduced to express these proteins, can be utilized for detection of the corresponding receptors and ligands, screening of novel low-molecular pharmaceuticals, and so on.

The present invention also provides genes corresponding to the polynucleotide sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which cDNA polynucleotide sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited to coding sequences, 5' and 3' [illegible] regions, alternatively spliced exons, introns, [illegible]

The corresponding genes can [illegible] known methods using the sequence information [illegible] herein. Such methods include the preparation [illegible]

primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that has been separated from the adjacent coding sequences, if any, present in the genome of the organism from which the gene was isolated.

Organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein are provided. The desired change in gene expression can be achieved through the use of antisense polynucleotides or ribozymes that bind and/or cleave the mRNA transcribed from the gene (Albert and Morris, 1994, Trends Pharmacol. Sci. 15(7): 250-254; Lavarosky et al., 1997, Biochem. Mol. Med. 62(1): 11-22; and Hampel, 1998, Prog. Nucleic Acid Res. Mol. Biol. 58: 1-39; all of which are incorporated by reference herein). Transgenic animals that have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided. Transgenic animals that have modified genetic control regions that increase or reduce gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein). In addition, organisms are provided in which the gene(s) corresponding to the [illegible] disclosed herein have been [illegible] partially or completely [illegible] the corresponding gene(s) [illegible] extraneous sequences into the corresponding gene(s) through deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished

through insertion, preferably followed by imprecise excision, of transposable elements (Plasterk, 1992, Bioessays 14(9): 629-633; Zwaal et al., 1993, Proc. Natl. Acad. Sci. USA 90(16): 7431-7435; Clark et al., 1994, Proc. Natl. Acad. Sci. USA 91(2): 719-722; all of which are incorporated by reference herein), or through homologous recombination, preferably detected by positive/negative genetic selection strategies (Mansour et al., 1988, Nature 336: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153; 5,614, 396; 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably are mammals. Such organisms are useful for the development of non-human models for the study of disorders involving the corresponding gene(s), and for the development of assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s). Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known [illegible]iques for determination of such domains from sequence [illegible]

Proteins and [illegible] [illegible]ention include proteins with amino acid sequence [illegible] at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more

preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologs of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide, as determined by those of skill in the art. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous, [illegible]

The invention also includes [illegible] sequences complementary to those of the polynucleotides disclosed herein.

The present invention also includes polynucleotides

capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table 33 below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

Table 33

| Stringency Condition | Polynucleotide Hybrid | Hybrid Length (bp)‡ | Hybridization Temperature and Buffer† | Wash Temperature and Buffer† |
|---|---|---|---|---|
| A | DNA : DNA | ≥50 | 65℃; 1×SSC -or- 42℃; 1×SSC,50% formamide | 65℃; 0.3×SSC |
| B | DNA : DNA | <50 | $T_B$*; 1×SSC | $T_B$*; 1×SSC |
| C | DNA : RNA | ≥50 | 67℃; 1×SSC -or- 45℃; 1×SSC,50% formamide | 67℃; 0.3×SSC |
| D | DNA : RNA | <50 | $T_D$*; 1×SSC | $T_D$*; 1×SSC |
| E | RNA : RNA | ≥50 | 70℃; 1×SSC -or- 50℃; 1×SSC,50% formamide | 70℃; 0.3×SSC |
| F | RNA : RNA | <50 | $T_F$*; 1×SSC | $T_F$*; 1×SSC |
| G | DNA : DNA | ≥50 | 65℃; 4×SSC -or- 42℃; 4×SSC,50% formamide | 65℃; 1×SSC |
| H | DNA : DNA | <50 | $T_H$*; 4×SSC | $T_H$*; 4×SSC |
| I | DNA : RNA | ≥50 | 67℃; 4×SSC -or- 45℃; 4×SSC,50% formamide | 67℃; 1×SSC |
| J | DNA : RNA | <50 | $T_J$*; 4×SSC | $T_J$*; 4×SSC |
| K | RNA : RNA | ≥50 | 70℃; 4×SSC -or- 50℃; 4×SSC,50% formamide | 67℃; 1×SSC |
| L | RNA : RNA | <50 | $T_L$*; 2×SSC | $T_L$*; 2×SSC |
| M | DNA : DNA | ≥50 | 50℃; 4×SSC -or- 40℃; 6×SSC,50% formamide | 50℃; 2×SSC |
| N | DNA : DNA | <50 | $T_N$*; 6×SSC | $T_N$*; 6×SSC |
| O | DNA : RNA | ≥50 | 55℃; 4×SSC -or- 42℃; 6×SSC,50% formamide | 55℃; 2×SSC |
| P | DNA : RNA | <50 | $T_P$*; 6×SSC | $T_P$*; 6×SSC |
| Q | RNA : RNA | ≥50 | 60℃; 4×SSC -or- 45℃; 6×SSC,50% formamide | 60℃; 2×SSC |
| R | RNA : RNA | <50 | $T_R$*; 4×SSC | $T_R$*; 4×SSC |

‡: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, [illegible] of the polynucleotides [illegible] regions [illegible]

† SSPE (1×SSPE is 0.15M NaCl, 10mM $NaH_2PO_4$, and [illegible] can be substituted for SSC (1×SSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

*$T_B$ - $T_R$: The hybridization temperature for hybrids anticipated to be less than

50 base pairs in length should be 5-10°C less than the melting temperature ($T_m$) of the hybrid, where $T_m$ is determined according to the following equations. For hybrids less than 18 base pairs in length, $T_m$(°C)=2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, $T_m$(°C)=81.5 + 16.6($\log_{10}$[$Na^+$]) + 0.41 (%G+C) - (600/N), where N is the number of bases in the hybrid, and [$Na^+$] is the concentration of sodium ions in the hybridization buffer ([$Na^+$] for 1×SSC=0.165M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and Current Protocols in Molecular Biology, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

Preferably, each such hybridizing polynucleotide has a length that is at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.